

VR 空間内におけるアバタとの視線コミュニケーション

小宮山撰^{†1} 亀川真吾^{†1} 柿沼育^{†1} 盛川浩志^{†1†2}

概要: 本研究は、HMD 装着者が VR 空間内のアバタと対話する状況における視線検出の重要性を述べている。実験環境として対面販売を想定した VR 世界を構築し、対話が音声と視線停留により進行するようにスクリプトを作成した。被験者にアバタとの対話のタスクを実行させた結果、アイトラッキングを用いて視線を検出した被験者グループと、視線を顔の向きで代用した被験者グループにおいて、視線の停留する場所に違いが見られた。また、アイトラッキングを用いた方がアバタとの空間共有感およびアイコンタクト感について有意に評価が高いことが分かった。

キーワード: バーチャルリアリティ, コミュニケーション, アバタ, 視線検出, HMD

Eye communication with avatar in VR space

SETSU KOMIYAMA^{†1} SHINGO KAMEGAWA^{†1} IKUMU KAKINUMA^{†1}
HIROYUKI MORIKAWA^{†1†2}

Abstract: In this research, the importance of eye-tracking when a subject wearing HMD interacts with avatar in VR space is described. We constructed a VR world that assumed face to face sales as an experimental environment and created a script so that the conversation will progress by the subject's gaze and speech. As a result of letting the subject work in conversation with avatar, a difference was observed in the place where gaze stopped between the group of subjects whose gaze were detected by eye-tracking and one whose gaze were substituted with face direction. Moreover, space sharing feeling and eye contact feeling with avatar when using eye tracking were highly evaluated.

Keywords: Virtual Reality, Communication, Avatar, Eye Tracking, HMD

1. はじめに

対面コミュニケーションにおいて、視線は非言語情報として重要な役割を占めており、アイコンタクトによってお互いに発話タイミングを図り、話の内容により目を合わせたり、そらす行為を無意識のうちに行っている[1][2]。一方で、非対面コミュニケーションでは、非言語情報の伝達が不十分であるため、人間の視線や表情をアバタに反映させる、アバタを用いたコミュニケーションシステムの研究が精力的に行われている[3][4]。HMD を用いたアバタとの対話では視線の代わりに首の向きを用いた研究もある[5]。最近では VR 空間でのアバタを介した人対人のコミュニケーションのために、HMD に様々なセンサーを仕込み、装着者の表情や動作をアバタに反映させる研究が CG の研究分野で進んでいる[6][7]。心理学の研究分野でも実験パラメータの自由度を大きくできることから、アバタを用いた VR を積極的に利用する動きがある[8]。

一般的な HMD には表情はおろか視線の検出機能すら無いため、ほとんどの VR アプリケーションでは視線の情報が必要な場合は HMD 自体のトラッキング、すなわち首の向きで代用してきた。しかし、これまでの研究においても視線の代わりに首の向きを用いることの具体的な問題点に

ついては明確になっていない。アバタを介した人対人の対話ではなく、コンピュータのアバタとの対話であれば、首の向きでも実用上十分の可能性もある。そこで、本研究では近年製品化された視線検出機能を有する HMD を用い、HMD ユーザが VR 空間内のアバタとコミュニケーションをとるアプリケーションを想定した対話実験を行う。HMD ユーザが見ている場所をアイトラッキングで求める場合と、HMD の向きとする場合の 2 条件で、実際の視線方向とアバタに対する主観的評価を比較する。

2. 実験装置

2.1 概要

本実験では視線が合う、同じものを見る、という行為が自然に行われるように、HMD ユーザを購買者、アバタを販売員とする対面販売を模擬した VR システムを作成する。対面販売では相手を見る、特定の商品を見るという行為が指示されなくても自然に発生すると期待できる。

対話システムの VR 世界は Unity を用いて作成する。直方体の部屋の中に机が配置され、机を挟んで HMD ユーザと対面するようにアバタを配置する。机の上には商品と見立てた 3 種類のスマートフォンが置かれている。

被験者は HMD を装着し、着席した体勢で VR 空間を観察する。同時に被験者はマイク入力可能なイヤホンを装着し、アバタと対話する。被験者が注視する対象に応じてア

^{†1} 青山学院大学
Aoyama Gakuin University
^{†2} 現所属：早稲田大学
Waseda University

バタが発声し、アバタからの質問に対する被験者の「はい」、「いいえ」の応答により場面が進行していく。被験者の注視対象の判定は「視線方向」あるいは「首の向き」のいずれかで行う。会話中の被験者の視線方向および首の向きを連続的に取得し、タスク終了後にアンケートによる主観評価を行う。

2.2 HMD[9]

実験に使用する HMD は視線追跡機能を有する FOVE を用いる。FOVE は赤外線アイトラッキング技術によりユーザの視線を追跡することが可能である。これは一般的な HMD である HTC Vive や Oculus Rift には無い機能である。また、頭部のトラッキングシステムとして、オリエンテーショントラッキング IMU (Inertial Measurement Unit) と、赤外線ポジショントラッキングを搭載している。オリエンテーショントラッキングは頭の傾きを検知し、ポジショントラッキングは外付けカメラと赤外線システムを利用し、位置の検出が可能である。仕様によれば視線のトラッキング精度は 1 度以下である。

2.3 音声

本実験では、被験者とアバタは音声を用いて対話する。アバタの音声は音声合成とリップシンクアセットを導入し、アバタは口を動かしながら話しているように見える。被験者の発話音声の認識については、IntelRealSenseSDK により、マイク入力を文字列に変換させている。ただし、本実験では非言語情報によるコミュニケーションを主たる検討対象としているため、誤認識により会話の自然さが損なわれないように、ユーザの返答は「はい」と「いいえ」のみに限定させ、その他の言葉は、返答に反映されない。

2.4 アバタの視線と動作

2.4.1 視線の動き

使用するアバタはニコニコ立体が無償で提供している「プロ生ちゃん MMD Ver.3 モデル」を使用する。このキャラクターモデルは MikuMikuDance 形式であるため、MMD4Mecanim を用いて Unity 上で使用可能な形式に変換し用いている。

アバタの視線および手の動作の例を図 2.1 に示す。アバタの視線の動きは大きく二つあり、一つは被験者を見つめる動きである。被験者が頭を左右に動かすと、アバタの視線もユーザの動きに合わせて追従する。もう一つはスマートフォンを見る動きである。いずれの場合もアバタは視線のみ動き、アバタの首は視線方向に動かない。

アバタの視線を動かすパターンとして会話中とそれ以外の場合に分けられる。会話中では、被験者に返答を求める会話では被験者を見続け、スマートフォンの説明中は被験者とそのスマートフォンを 2 秒間ごと交互に見る動きに



図 2.1 VR 空間のオブジェクト配置とアバタの挙動
上から、A) 被験者を見るアバタ、B) 左のスマートフォンを見るアバタ、C) 被験者が見ているスマートフォンを見るアバタ、D) スマートフォンを指し示し説明するアバタ。

なっている。それ以外の場合は、ユーザとランダムに選ばれたスマートフォンを 3 秒間ごと交互に見る動きになっている。その際に、ユーザが 20 秒以上スマートフォンを選択しない状態のときは、「何か気になるスマートフォンを見てください」と発話し注意を促す。またいずれの場合も、ユ

ーザに返答を求める会話を除き、ユーザのしているオブジェクトにアバタは優先的に視線を合わせる。なお、スマートフォンを見るアバタの眼球の向きは、事前に自然に見えるように手動で設定する。

2.4.2 ジェスチャー

アバタの動作アニメーションは最小限の動作のみ実装している。スマートフォンの説明を開始する際に、「選択されたスマートフォンに近づき、手で指し示す」という動作である。ただし、アバタのアニメーションは目の動きとは独立なので、アニメーション中であってもユーザのしている箇所によってアバタの視線も動く。

2.5 被験者の視線と首の向きの検出

本装置は、毎フレームごとの時間、HMDのRotationの角度、視線のPosition、首振りのPosition、視線によって当たり判定されたオブジェクト、首の向きによって当たり判定されたオブジェクトを書き出す。当たり判定とは、UnityのRaycast機能[10]を用いてRay（光線）を飛ばしRayが衝突したオブジェクトの情報を取得するもので、本実験でRayの衝突判定を行うオブジェクトは、アバタの顔、机にある3つのスマートフォンである。それ以外はNoneとしている。Rayが2秒以上オブジェクトのBoxColliderに停留すると、オブジェクトを選択したという判定になる。

なお、被験者側のデータと平行してアバタの視線方向（ユーザ、3つのスマートフォン）も書き出す。

2.6 対話システムのフロー

次に対話システムの流れについて述べる。システムのフローチャートを図2.2に示す。初めに実験者が被験者に実験の説明を行い、被験者に視線検出のキャリブレーションを行わせる。対話は3つのフェーズからなる。各フェーズの初めにアバタが「気になるスマートフォンを見てください」と発声し、被験者がスマートフォンを2秒以上見ると、アバタがそれぞれのスマートフォンの説明を開始する。フェーズ1では名称および状態について、フェーズ2ではそれぞれのスマートフォンの値段と画面サイズについて、フェーズ3では、それぞれのスマートフォンのロック解除方法とおすすめの点についての説明を行う。なお、フェーズ1から、ユーザの視線データおよび首振りデータがファイルに書き出される。

被験者が音声で返答するタイミングは、被験者が特定のスマートフォンを見た際に、「このスマートフォンの説明を聞きますか」とアバタから問いかけられたとき、一つのフェーズですべてのスマートフォンの説明を聞き終えた際に、「もう一度説明を聞きますか」とアバタが問いかけられたときの2回である。

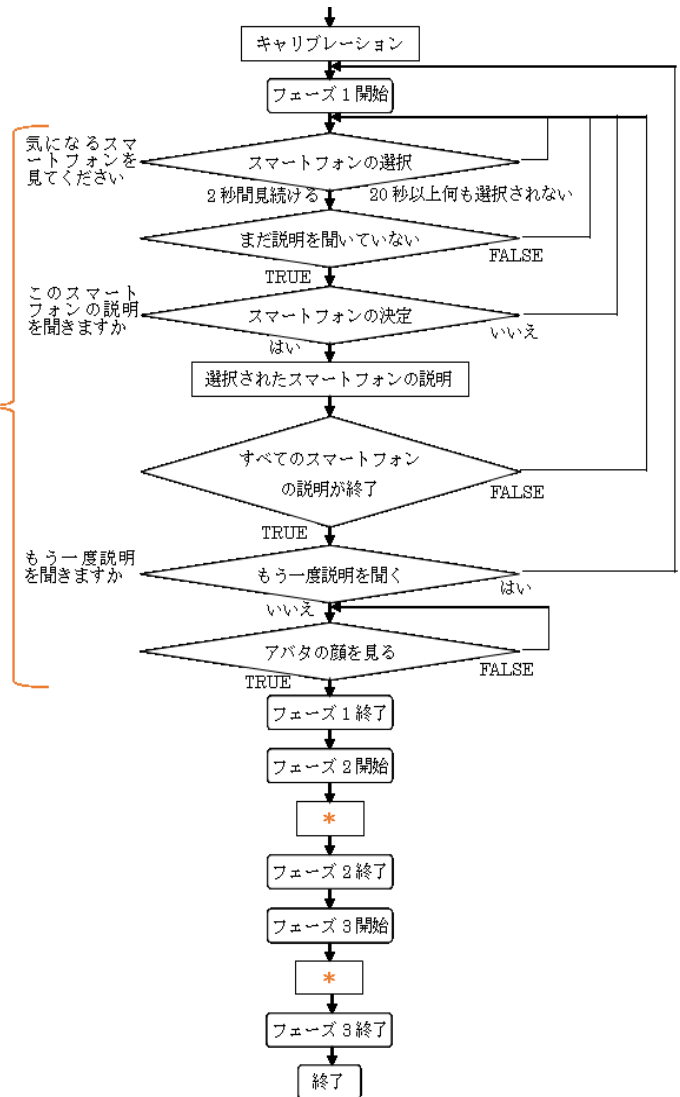


図 2.2 実験システムのフローチャート

3. 評価実験

3.1 実験方法

大学生および大学院生計 24 名（平均年齢 21.9 歳の男性 23 名、女性 1 名）を被験者としてアバタとの対話実験を行った。被験者の内訳は、HMD を頻繁に使用する人 10 名、数回使用したことがある人 6 名、一度も使用したことがない人 8 名である。また、本実験では当たり判定を視線で行う被験者 12 名と首振りで行う被験者 12 名に分けて実験を行った。なお、被験者にはどちらの当たり判定で実験を行うかは知らせていない。

実験の様子を図 3.1 に示す。被験者には、カメラと HMD の距離をなるべく保つように指定された位置にある椅子に座るよう指示した。被験者に 2.6 節で示したフローを実施させ、実験中の被験者の視線および首振りの方向を取得し、終了後にアンケートによる主観評価を行った。

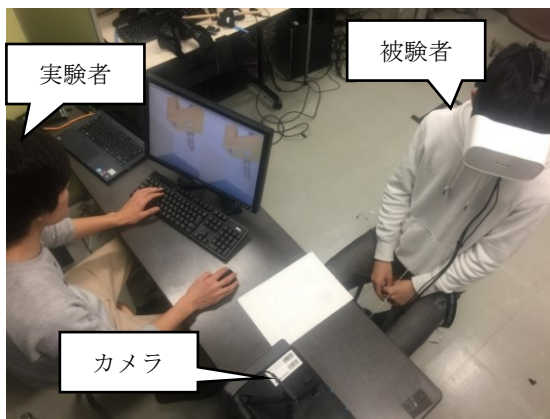


図 3.1 実験の様子

3.2 主観評価

アンケートによる評価項目を表 3.1 に示す。

表 3.1 主観評価項目

- | | |
|---|---------------------|
| A | アバタに親しみが持てたか |
| B | アバタと同じ空間にいると感じたか |
| C | アバタと視線が合っていると感じたか |
| D | アバタがどこを見ているのか理解できたか |
| E | アバタと同じ対象物を見ていると感じたか |
| F | アバタの視線に違和感を覚えなかったか |
| G | 発話のタイミングが取れやすかったか |
| H | 対象物の選択にストレスを感じなかったか |
| I | 目に疲れを感じなかったか |
| J | 首に疲れを感じなかったか |

各項目を選定した理由を次に挙げる。A, B, F の項目は本実験で使用したアバタについて、被験者がどのように感じるかを総合的な評価を得るために設定した。コミュニケーションにおいて親しみが持てること、同じ空間にいることはよいコミュニケーションにつながることを考慮した。また本実験ではアバタが視線を動かす際に、被験者はアバタの首の動きから視線を判断してしまうことを避けるため、首の動きはしないものとした。その際に違和感を覚えるかについて聞くことで、今後のシステム開発につながると考えた。C, D, E の項目は VR 空間における視線コミュニケーションの総合的な評価を得るために設定した。視線コミュニケーションにおいて、視線が合う、視線がそれることを C, D の評価項目によって得ることができ、また E を設定したのは、VR 空間ならではの同じ対象物を見るという視線コミュニケーションを考えたからである。G の項目は、アバタとのコミュニケーションに視線を用いることによって HCI における視線の役割の一つである、話す・聞くの交換時期の調整の評価を得るために設定した。H, I, J の項目は対象物の選択に視線を用いる場合は、首を振らずに選択することが可能であるが、視線操作によって目をより動

かすので、目と首に対する疲労度に影響するのではないかと考えたからである。

評価項目は点数がそれぞれの項目において、-3 点から 3 点の 7 段階評価で行い、点数が高いほど、よい評価となるように設定した。また、対話実験タスク中に被験者がアバタを見た回数、アバタが問いかけたときの返答時に被験者の視線と首振りに向いている対象物を客観データから抽出した。

4. 実験結果

4.1 タスク完了時間

当たり判定が視線の場合と首振りの場合におけるキャリアレーション時間を除いたタスク完了時間の平均を比較した棒グラフを図 4.1 に示す。

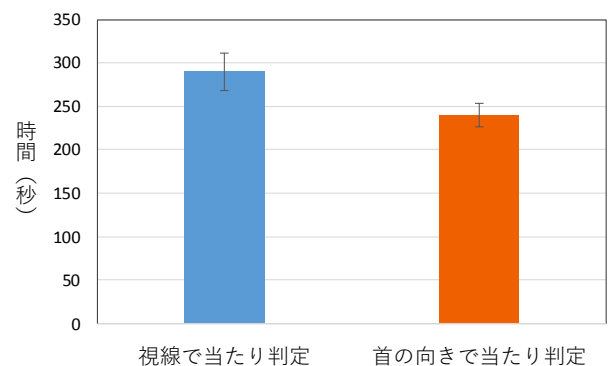


図 4.1 2つの被験者グループにおけるタスク完了時間

4.2 対話実験タスク中にアバタを見た回数

実験全体で被験者の視線と首の向きがアバタの方に向いた回数を図 4.2 に示す。なお、視線と首の向きが同時にアバタに向くときは首の向きのみカウントし、視線で見た回数には含めていない。

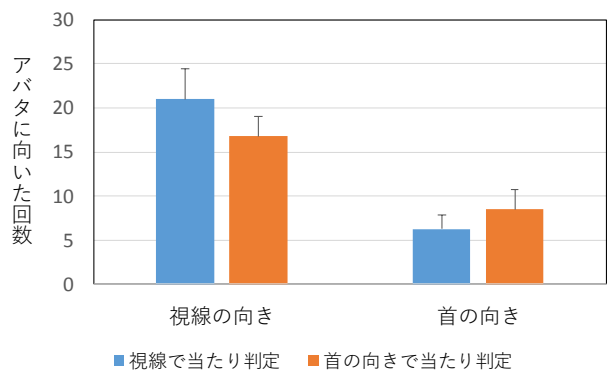


図 4.2 被験者の視線と首の向きがアバタに向いた回数

有意差水準を 5% とした分散分析の結果、実験全体としては視線と首がアバタに向いた回数に関して視線による当たり判定条件と首の向きによる当たり判定条件の間に有意差は見られなかった。

4.3 返答時に視線と首振りが向いている対象物

人対人の会話では返答する際に相手の顔を見ることは重要な要素である。そこでアバタからの質問に返答するタイミングにおける被験者の視線と首の向きが向いているオブジェクトをカウントした。視線の結果を図 4.3 に首の向きの結果を図 4.4 に示す。図中で*は危険率 5%で有意差があることを示す。

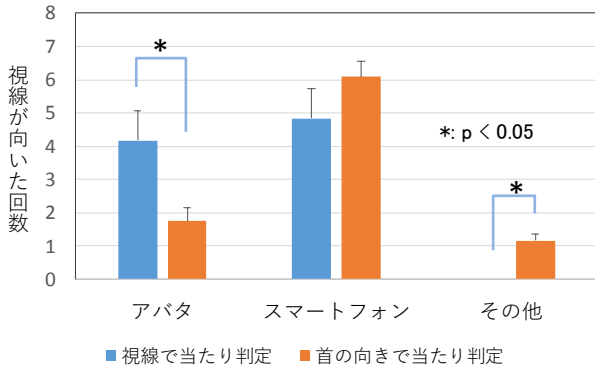


図 4.3. 返答時に被験者の視線が向くオブジェクト

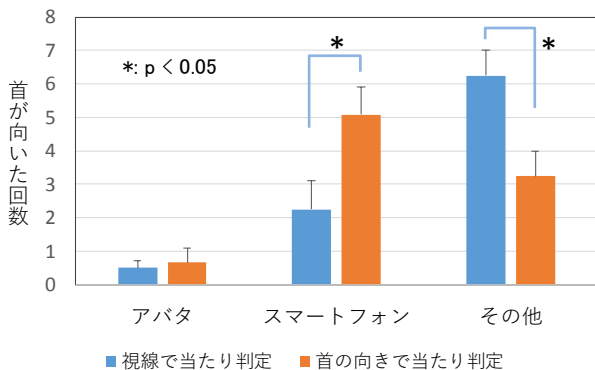


図 4.4. 返答時に被験者の首が向いているオブジェクト

4.4 主観評価結果

アンケート結果の 10 の評価項目を図 4.5 に示す。全般的に当たり判定に視線を用いる方が高い評価となったが、有意水準を 5%とした分散分析の結果、有意差が認められた項目は、評価項目 B の「アバタと同じ空間にいると感じたか」と、評価項目 C の「アバタと視線が合っていると感じたか」であった。また、どちらの当たり判定でも否定的な評価となった項目は評価項目 F の「アバタの視線に違和感を覚えなかったか」が唯一であった。評価項目 H の「対象物の選択にストレスを覚えなかったか」は 5%有意にはならなかったが、首の向きで当たり判定した被験者グループにおいて、3 のプラス評価を与える者が 3 人いた一方で-3 のマイナス評価を与える者が 2 人存在しており、非常にばらつきが大きくなった。

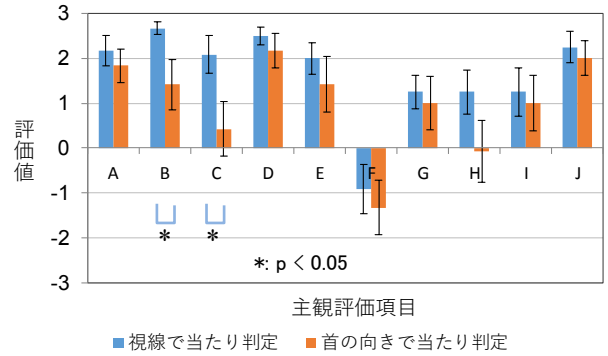


図 4.5. 主観評価結果

5. 考察

5.1 アバタを見た回数に関する考察

対話実験タスク中に被験者の視線と首振りがアバタの方向に停留する回数には、当たり判定の違いによる有意差は生じなかった。被験者は事前にどちらの当たり判定で実験タスクを行うかを知らされないため、対象物を選択する手法の差異がアバタを見る回数に大きな影響を与えないのは当然といえる。

首振りよりも視線のみでアバタを見た回数の方が多くなった要因としては、被験者は首を動かさずに目の動きのみで見ることができ、身体的負担が少ないためと考えられる。本実験のように HMD を装着して VR 空間のアバタと向かい合う状況であっても、視線のみでアバタを見る動作が多くなるといえる。

5.2 返答時に見る対象物に関する考察

アバタの問いかけに対する返答時に被験者の視線と首振りが向いている対象物は、図 4.3 と図 4.4 からわかるように、当たり判定が視線のときは、視線はアバタとスマートフォンに同程度向き、首はアバタとスマートフォン以外に向く傾向がある。これは、被験者がアバタとスマートフォンの中間に首を向け、目の動きでアバタやスマートフォンを見ているためと考えられる。

当たり判定が首振りのときは、被験者の視線の向いている対象物はスマートフォンが最も回数が多く、アバタに向いている回数は少ない。また、首もスマートフォンに向いている回数が一番多い。当たり判定が首の向きであるため、首をスマートフォンに向けざるをえず、改めてアバタの顔を見るためには首を回す必要が出てくる。そのため、負担を感じ、視線がスマートフォンに向き続けたと考えられる。

5.3 主観評価結果に関する考察

評価項目 B の「アバタと同じ空間にいると感じたか」について、当たり判定が首振りに比べて、視線のほうがよりアバタと同じ空間にいると感じたという結果が出た。実験

タスク中にアバタを見た回数に有意差はないため、アバタからの問いかけに対する返答時にアバタを見ること、「アバタと同じ空間にいる感覚」に寄与したと考えられる。また被験者の感想に、「アバタの目が動くと一緒にいる感じがする」というものがあり、当たり判定が視線の方がアバタを視野に入れやすく、目の動きを確認しやすかったと思われる。

評価項目 C の「アバタと視線が合っていると感じたか」についても、当たり判定が視線の方がよりアバタと視線が合っていると感じたという結果が出た。これも返答時において、アバタに視線を向ける頻度が高いことが、「アバタと視線が合っている感覚」に影響したと考えられる。

評価項目 F の「アバタの視線に違和感を覚えなかったか」については、否定的な結果となった。これは、本実験で使用したアバタの視線が、首の動きと連動していないため不自然さが出てしまったと考えられる。被験者の感想でも、「同じ対象物を見るときのアバタの視線の距離感が合っていない」という意見があり、アバタは説明時に少し前に出るため、視線がずれて見える恐れがある。首の動きや身体運動を伴う際の目の動きは細かな調整が必要といえる。

5.4 全体の考察

実験タスク中にアバタを見た回数は、全体としては当たり判定が視線の場合と首振りの場合で大きな差は生じないにもかかわらず、アバタの問いかけに対する返答時に差が生じた。対話を進行させるために被験者はスマートフォンを必ず見る必要があるが、その際に首を向ける必要があるため、その後の動作が変わることを示している。今回の実験では被験者はスマートフォンを見ている最中にアバタから問いかけられ、アバタに返答することが求められる。この際にアバタを見る行為はその後の対話進行に必要な動作ではないため、当たり判定が首振りで行われた場合には、労を惜しんで省略されやすいと考えられる。しかし、現実の会話では返答時に相手を見ることは自然であり、相手との距離感を測る重要な行為である。VR 空間においてアバタとのコミュニケーションを行う場合、ユーザの視線を首の向きで代用することは自由な視線の動きに影響を与え、結果的にアバタと視線が合う感覚や空間の共有感覚を阻害する恐れがあるといえよう。

一方、視線を利用することには問題点もある。アイトラッキング可能な FOVE はキャリブレーションを行った装着位置から HMD がずれないように注意する必要がある。アイトラッキング機能のない HMD ではオブジェクトの選択は HMD の視野の正面方向にマーカーを表示し、そのマーカーをオブジェクトに重ねるといった方法が一般的であるため、装着後に HMD が多少ずれてしまっても再調整する必要があまりない。従って単にオブジェクトを選択する目的であれば視線を用いるメリットはあまりないといえる。

6. 結論

本研究では、VR 空間での対象物の選択に視線を向ける手法として、アイトラッキング技術を用いた場合と、被験者の首の向きを視線として用いた場合を比較することで、VR 空間におけるアバタとのノンバーバルコミュニケーションに起きる違いを検討した。

対象物の選択にアイトラッキング技術を用いた場合と、被験者の首の向きを視線として用いた場合では、実験全体としてアバタを見た回数に大きな差はなかった。しかし、首の向きを当たり判定として利用する方法では、アバタから問いかけられた際の返答時に被験者がアバタを見る頻度が低くなることが分かった。主観的にも「アバタと同じ空間にいる感覚」および「アバタと視線が合っている感覚」の二項目において評価が低くなった。従って、VR 空間においてアバタとの空間共有感やアイコンタクト感を高めるためには、被験者の視線を検出してコミュニケーションの進行を制御することが重要である。首の向きを視線の代用とした場合は HMD ユーザが不自然さを感じる恐れがある。

VR 空間においてアバタと対話する場合、HMD ユーザ自身も VR 空間の一員であり、HMD ユーザの非言語情報をシステムが取得する必要があるが、アバタとの自然な対話を実現するためには、最低限 HMD ユーザの視線の検出は必要であると言える。

参考文献

- [1] 武川直樹：コミュニケーションにおける視線の役割，電子情報通信学会誌，Vol.85, No.10, pp756-760, 2002
- [2] 王翼，宇津木成介：対人コミュニケーションにおける視線の役割---視線と対人距離の関係，感情心理学研究，Vol.18, No.3, pp188, 2011.
- [3] Kunal Gupta, Gun A. Lee, Mark Billinghurst, Do You See What I See? The Effect of Gaze Tracking on Task Space Remote Collaboration, IEEE Trans. on Visualization and Computer Graphics Vol. 22, No. 11, Nov. 2016
- [4] 高林範子，山本真代，小野光貴，渡辺富夫，石井裕：アバタに微笑みと眼球動作モデルを付加した看護コミュニケーション教育支援システム，人間工学，Vol.52, No.3, 2016.
- [5] 根岸匠，神場知成，田中二郎：バーチャルキャラクターとの視線によるコミュニケーションを実現する添い寝システムの開発，情報処理学会インタラクション2014論文集，pp624-628, 2014.
- [6] Hao Li, Laura Trutoiu, Kyle Olszewski, Lingyu Wei, Tristan Trutna, Pei-Lun Hsieh, Aaron Nicholls, Chongyang Ma, Facial Performance Sensing Head-Mounted Display, ACM Trans. on Graphics, Vol. 34, No. 4, Article 47, Aug. 2015
- [7] Guillermo Bernal, Pattie Maes, Emotional Beasts: Visually Expressing Emotions through Avatars in VR, CHI 2017, Denver, CO, USA, May 6-11, 2017
- [8] Xueni Pan, and Antonia F. de C. Hamilton, Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape, British Journal of Psychology, 2018
- [9] FOVE の概要：https://www.getfove.com
- [10] Raycast の概要：https://docs.unity3d.com/jp/540/ScriptReference/Physics.Raycast.html