

VR 空間操作コマンドとしてのアイジェスチャ UI 特性分析 (1) ～単一アイジェスチャの UI 特性分析～

夏目達也^{†1} 内村裕也^{†2} 柴田史久^{†1} 木村朝子^{†1}

概要：VR 技術の発達に伴い、身体の動きを入力方法とした手法の研究が活発になっている。最近ではアイジェスチャと呼ばれる視線移動のような目の動きを入力とする研究も登場してきた。しかし、VR 空間においてアイジェスチャを入力として用いた研究はまだ数が少なく、知見は十分ではないと考えられる。そこで本研究では、各アイジェスチャ（注視、瞬き、ウインク、見開く、細める）に対して、VR 空間における入力操作を 4 種類に分類したものをタスクとして設定した実験（選択、掴み、色変更、拡大）を実施した。その結果を基に、各アイジェスチャにはどのような特性があり、どのような操作に適しているのかについての分析・考察を行った。

キーワード：HMD, VR, アイジェスチャ, 視線入力, ユーザインタフェース

1. はじめに

ポスト WIMP 研究が活発化する中で、Microsoft 社の Kinect に見られるような、入力として身体部位を利用するユーザインタフェース (UI) が増えてきている。一般的に言えば「物を見るために目を開く」「眠るために目を閉じる」「移動するために歩く」のように、人は目的のために自然に身体を動かすことを身につけている。情報機器に対する恣意的な入力操作でも、自分の身体機能の 1 つとして自然に扱えることが望ましい。

身体部位を利用する UI には、これまでハンドジェスチャといったユーザの手や腕の動きを入力として利用するものが多かったが[1][2]、アイトラッキング機能の搭載された HMD が登場し、VR や AR/MR 空間への入力手段として瞬きや視線移動といった目の動き（以降、アイジェスチャ）を入力として利用する UI も増えてきている[3]。

こうした「アイジェスチャ」は、他の身体動作やスイッチやボタン等の器具で代替できないかといえ、できなくもない。しかし、片手/両手が別の用途やデバイスで塞がっている場合は、ハンズフリーで使える「アイジェスチャ技術」は有力な操作ツールとなる。特に、VR や AR/MR のような空間体験の場合は、両手両足を扱うことも多く、「アイジェスチャ UI」の存在意義が高い。

例えば、Rajanna[4]らは、VR 空間におけるキーボード操作方法として、視線移動と注視時間を組み合わせた手法と、視線移動とコントローラのボタン入力を組み合わせた手法を提案、比較している。しかし、VR や AR/MR 空間におけるアイジェスチャ UI の研究は、まだ数が少なく、知見も十分ではない。

そこで本研究では、VR 空間操作において、各種アイジェスチャにどのような特性があるのか、そしてどのようなアイジェスチャがどのような操作に適しているのかについ

て分析する。

2. アイジェスチャの検討

2.1 アイジェスチャの整理

まずは制約条件なしで、どのようなアイジェスチャがあり得るから考えはじめた。

30 名 (21 歳～33 歳の男性 26 名、女性 4 名) の人に思いつくアイジェスチャを全て列挙してもらったところ、「ウインク」「瞬き」「視線の動き (例、上から下に視線移動)」「注視」「両目を閉じる」「片目を閉じる」「目を細める」「寄り目」「目を見開く」「白目をむく」といったアイジェスチャが挙げられた。

これらのアイジェスチャの中から、本稿ではまず比較的 UI として利用しやすい「注視」「瞬き」「ウインク」「見開く」「細める」の 5 種類のアイジェスチャをピックアップした。

2.2 アイジェスチャによる VR 空間操作の分類

【単発的入力と継続的入力】

アイジェスチャを操作コマンドとして利用する際、例えば同じ「目を閉じる」アイジェスチャでも、目を閉じるたびに値が変わるような使い方と、目を閉じている間値が変わり続けるといった使い方があり得る。本研究では、前者を「単発的な入力」、後者を「継続的な入力」と呼ぶ。

単発的な入力とは、ボタンを押す動作やマウスのクリックのような、入力動作を行う毎に入力が行われる動作のことであり、アイジェスチャを行うたびに入力が実行されるといった入力方法のことである。

一方、継続的な入力とは、指を押し広げるピンチ操作やマウスのスクロール操作のような入力動作を行っている間は入力が継続される動作のことであり、アイジェスチャを行っている間は入力が継続されるといった入力方法である。

【ON/OFF 操作と値変更操作】

^{†1} 立命館大学 情報理工学部
School of Information Science and Engineering, Ritsumeikan University

^{†2} 現在、NTT データ
NTT DATA Corporation

さらに本研究では、アイジェスチャによる入力の結果、ON/OFFのように2種類の状態が切り替わる場合と、それよりも多い3種類以上の状態が切り替わる場合で区別することとした。

これらを踏まえて、本研究ではアイジェスチャで行うVR空間操作を、単発的な入力であるか、継続的な入力であるかという2種類の入力と、入力の結果ON/OFF操作が行われるか値変更操作が行われるかという2種類の操作を組み合わせた4種類に分類する。本研究では、(A) 単発的な入力&ON/OFF操作、(B) 継続的な入力&ON/OFF操作、(C) 単発的な入力&値変更操作、(D) 継続的な入力&値変更操作の4種類の入力操作を対象に、2.1で述べた5種類のアイジェスチャの特性を評価する。

3. 実験準備

3.1 システム構成

実験で使用するシステムの構成を図1に、筋電位計測装置を図2に示す。実験では、赤外線によるアイトラッキング機能を有するHMDであるHTC VIVE Pro Eye（以降、VIVE Pro Eye）と筋電位計測装置（前頭筋に電極を貼付）を利用してアイジェスチャを認識する。実験タスクの構築には、Unityを用いている。

3.2 アイジェスチャ認識

「注視」「瞬き」「ウインク」「見開く」「細める」の5つのジェスチャを識別するために、「見開く」ではVIVE Pro Eyeと筋電位計測装置を併用し、それ以外のアイジェスチャではVIVE Pro Eyeのみを使って識別している。各々のアイジェスチャUIは、表1のようなパラメータを設定して認識している。

【注視】一定時間の間、対象を見続けると入力されるアイジェスチャである。Choeら[5]の研究では、注視を入力方法とした選択タスクにおいて、注視の時間が1.0秒、1.5秒、2.0秒の場合で比較しており、その結果1.0秒が最も評価が高いという結果を示している。そこで本研究の注視の設定時間も1000msとした。

【瞬き、ウインク】両目または片目を閉じて、また開くと入力されるアイジェスチャである。目を閉じてから開くまでの時間がある程度短ければ単発的な入力、長ければ継続的な入力と認識される。Kuら[6]の研究では、スマートフォン

で瞬きによる入力が既存のタッチ入力と比べてどの程度使いやすいのか比較する実験を行っている。その際、ウインクとして認識される時間を500msから1000msの間としていた。本研究では、これを参考に、瞬き、ウインクとして認識される目を閉じてから開くまでの時間を500msから1000msとした。

【見開く、細める】目を大きく見開いたり、細めたりするアイジェスチャである。これらのアイジェスチャを識別するために、目の開き具合の最大を1.0、最小を0.0としたときに、1.0を見開く、0.5以下0.1以上を細めるとして設定した。また、見開くを認識するにあたり、目の開き具合だけでは精度が不十分であったため、筋電位による判定も併用した。

3.3 評価指標

タスク達成までの平均作業時間、平均エラー回数、5項目の主観評価（表2）、そして好みのアイジェスチャの順位の結果から各アイジェスチャを評価・分析する。主観評価は非常に悪いを「1」、非常に良いを「7」とした7段階のリッカート尺度で評価する。



図1 システム構成

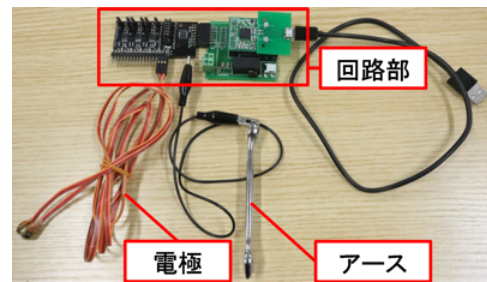


図2 筋電位計測デバイス

表1 アイジェスチャのパラメータ設定

| アイジェスチャ | 単発的なアイジェスチャ | 継続的なアイジェスチャ |
|---------|----------------------------|--------------------------------------|
| 注視 | 1000msの間、対象を見続けると入力 | 1000msの間対象を見続けることで入力が開始し、注視をやめると入力終了 |
| 瞬き | 500ms~1000msの間に両目を閉じて開くと入力 | 両目を閉じることで入力が開始、目を開くと入力終了 |
| ウインク | 500ms~1000msの間に片目を閉じて開くと入力 | 片目を閉じることで入力が開始、目を開くと入力終了 |
| 見開く | 目の開き具合が1.0であれば入力 | 目の開き具合が1.0である間、入力が継続 |
| 細める | 目の開き具合が0.1以上0.5以下であれば入力 | 目の開き具合が0.1以上0.5以下である間、入力が継続 |

4. 実験

4.1 目的

VR 空間操作コマンドとしての各アイジェスチャがどのようなタスクに適しているのか、そして各々にどのような UI としての特性が考えられるのかを分析するために、2.2 で挙げた(A)~(D)の 4 種類の入力操作ごとにタスクを設定し、実験を行う。

4.2 課題と手続き

4 種類の入力操作ごとでの実験タスクは以下の通り。

タスク 1: 単発的入力 & ON/OFF 操作 (選択操作)

VR 空間において、実験協力者の前方に 1~10 の数字が描かれた立方体をランダムに 10 個配置し、目標となる数字が描かれた立方体を視線で捉え、アイジェスチャを行うことで選択させる (図 3)。また、目標となる数字以外が描かれている立方体を誤って選択した場合、それはエラーとしてカウントされ、もう一度同じタスクを行わせる。

タスク 2: 継続的入力 & ON/OFF 操作 (掴み操作)

VR 空間において、実験協力者の正面に立方体を配置し、

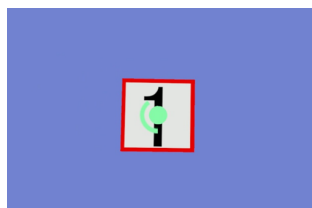
アイジェスチャを行うことで立方体を掴ませて目的地点まで視線によって移動させる (図 4)。目標地点に到達するまでに誤って掴みをやめてしまった場合や一度目標地点に到達したが掴み状態を解除する際に立方体が目標地点に接していなかった場合は、それらはエラーとしてカウントし、もう一度同じタスクを行わせる。

また、(B)のタスクでは入力方法として注視と瞬きは省いて実験を行う。これは、注視については、立方体から視線を外すと入力が終了されエラーとなってしまうタスクを完了できないためである。瞬きについては、両目を閉じ続けている間入力が行われるため、入力動作中は何も見えず、タスクを完了できないためである。

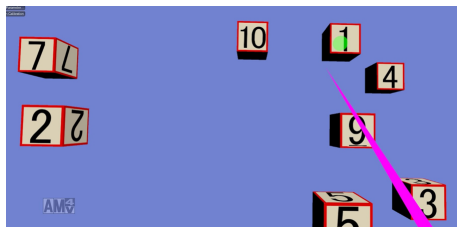
タスク 3: 単発的入力 & 値変更操作 (色変更操作)

VR 空間において、実験協力者の正面に立方体を配置し、アイジェスチャを行うことで立方体の色を目的の色になるまで変更させる (図 5)。また、誤って目的の色以外で決定した場合や目的の色で決定せず次の色に進んだ場合は、エラーとしてカウントされ、もう一度同じタスクを行わせる。

タスク 4: 継続的入力 & 値変更操作 (拡大操作)

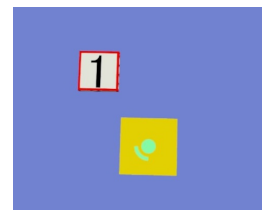


(a) 目標となる立方体の提示



(b) 実験シーン

図 3 単発的入力 & ON/OFF 操作 (選択操作) のタスク

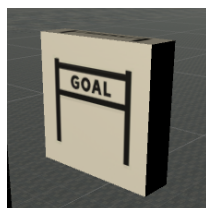


(a) 目標の色の提示

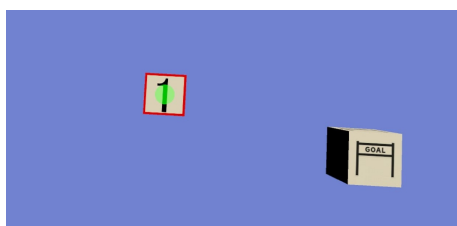


(b) 実験シーン

図 5 単発的入力 & 値変更操作 (色変更操作) のタスク

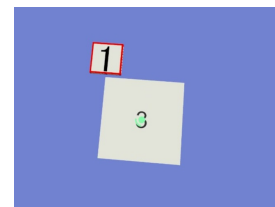


(a) 目的地点の提示

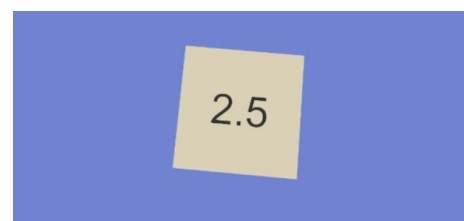


(b) 実験シーン

図 4 継続的入力 & ON/OFF 操作 (掴み操作) のタスク



(a) 目標となる立方体の大きさの提示



(b) 実験シーン

図 6 継続的入力 & 値変更操作 (拡大操作) のタスク

VR空間において、実験協力者の正面に立方体を配置し、アイジェスチャを行い続けることで立方体を目的の大きさになるまで拡大させる(図6)。元の立方体の大きさを1.0とし、目標となる立方体の大きさは3.0, 4.0, 5.0の3種類が用意されている。この3種類のうちの一つが目的の大きさとして目的の大きさとして設定される。また、目的の大きさの±0.2以内で拡大を停止できた場合は成功、目的の大きさの±0.2を超える場合はエラーとしてカウントし、もう一度同じタスクを行わせる。

また、タスク4では、入力方法として継続的なアイジェスチャの瞬きは省いて実験を行う。これは、タスク2と同様に継続的なアイジェスチャの瞬きは、両目を閉じ続けている間入力が行われるため、入力動作中は何も見えず、タスクを完了できないためである。

実験手順は以下の通りである。

- (1) 参加者毎に目の動きの検出精度に差が出ないように VIVE Pro Eye を用いてキャリブレーションを実行
- (2) 体験させる入力方法をランダムに決定し、入力方法の練習を行わせる
- (3) 実験タスクを行わせる
- (4) 手順(3)を同じ入力方法で3回行わせる
- (5) 体験した入力方法に対する主観評価を回答させる
- (6) 目の疲労が排除されるまで休憩を設ける
- (7) 残りの入力方法についても手順(2)~(6)を繰り返す
- (8) 体験した全ての入力方法を好みで順位付けさせる

4.3 条件

タスク1~タスク4の入力操作毎で使用するアイジェスチャは以下の通り。

タスク1とタスク3で使用するアイジェスチャは、単発的なアイジェスチャである注視、瞬き、ウインク、見開く、細めるの5種類。

タスク2で使用するアイジェスチャは、で評価する入力方法は、継続的なアイジェスチャであるウインク、見開く、細めるの3種類。

タスク4の実験で使用するアイジェスチャは、継続的なアイジェスチャである注視、ウインク、見開く、細めるの4種類。

4.4 参加者

実験協力者は20~22歳の男性12名で、矯正を含め、全員が正常視力を有した。

4.5 実験結果

タスク1~タスク4の入力操作に対する実験結果を、図7~図10に示す。平均作業時間のグラフでは、縦軸が1回の選択タスクを完了させるまでの平均作業時間、横軸は入力方法を表している。平均エラー回数の結果では、縦軸が1回成功するまでに失敗した回数、横軸は入力方法を表している。また、平均エラー回数の分析には、入力方法を要因として1要因分散分析を行った。主観評価の結果は、グ

表2 5項目の主観評価

| 評価項目 | 主観評価の内容 |
|------|--------------------------------|
| 1 | 容易さ(この入力方法をどれくらい簡単に扱えるか) |
| 2 | 学習容易性(この入力方法を覚えるのはどれくらい簡単だったか) |
| 3 | 適合性(タスクに対してこの入力方法はどれくらい適していたか) |
| 4 | 疲れにくさ(始める前と比べてどれくらい目が疲れなかったか) |
| 5 | 好み(全体的にこの入力方法をどれくらい好んだか) |

ラフの縦軸が7段階のリッカート尺度を表しており、横軸は使用したアイジェスチャの種類を表している。主観評価の分析には、入力方法を要因としてフリードマン検定を行った。入力方法の好みの順位の結果は、グラフの縦軸が入力方法、横軸が各順位を回答した人数を表している。好みの順位の分析には、入力方法を要因としたフリードマン検定とボンフェローニ法を用いた多重比較を行った。

以下、それぞれの入力操作毎での結果について述べる。

4.5.1 タスク1: 選択

単発の入力&ON/OFF操作では、容易さ(図7(c))において、入力方法間で有意差が認められた($p < 0.05$)。さらに、ボンフェローニ法を用いた多重比較を行った結果、注視とウインク、注視と見開く、注視と細めるの間に有意差が認められた($p < 0.05$)。また、学習容易性(図7(d))、疲れにくさ(図7(f))では入力方法の要因において有意差は認められたもの($p < 0.05$)、ボンフェローニ法を用いた多重比較を行った結果、入力方法間において有意差は認められなかった。好みの順位についても、入力方法間で有意差が認められた($p < 0.05$)。ボンフェローニ法を用いた多重比較を行った結果、注視と細めるの間に有意差が認められた($p < 0.05$)。

4.5.2 タスク2: 掴み

継続的な入力&ON/OFF操作では、平均エラー回数において、入力方法の要因で主効果が有意であった($F(2, 22) = 10.541, p < 0.05$)。下位検定として、入力方法間における有意差を確認するため、Ryanの方法による多重比較を行った。その結果、ウインクと細める、見開くと細めるの間に有意差が見られた($p < 0.05$)。主観評価では、容易さ(図8(c))において入力方法間で有意差が認められた($p < 0.05$)。さらに、ボンフェローニ法を用いた多重比較を行った結果、ウインクと細める、見開くと細めるの間に有意差が認められた($p < 0.05$)。同様に、学習容易性(図8(d))、適合性(図8(e))、好み(図8(g))に関してもフリードマン検定において入力方法間で有意差が認められ、ボンフェロ

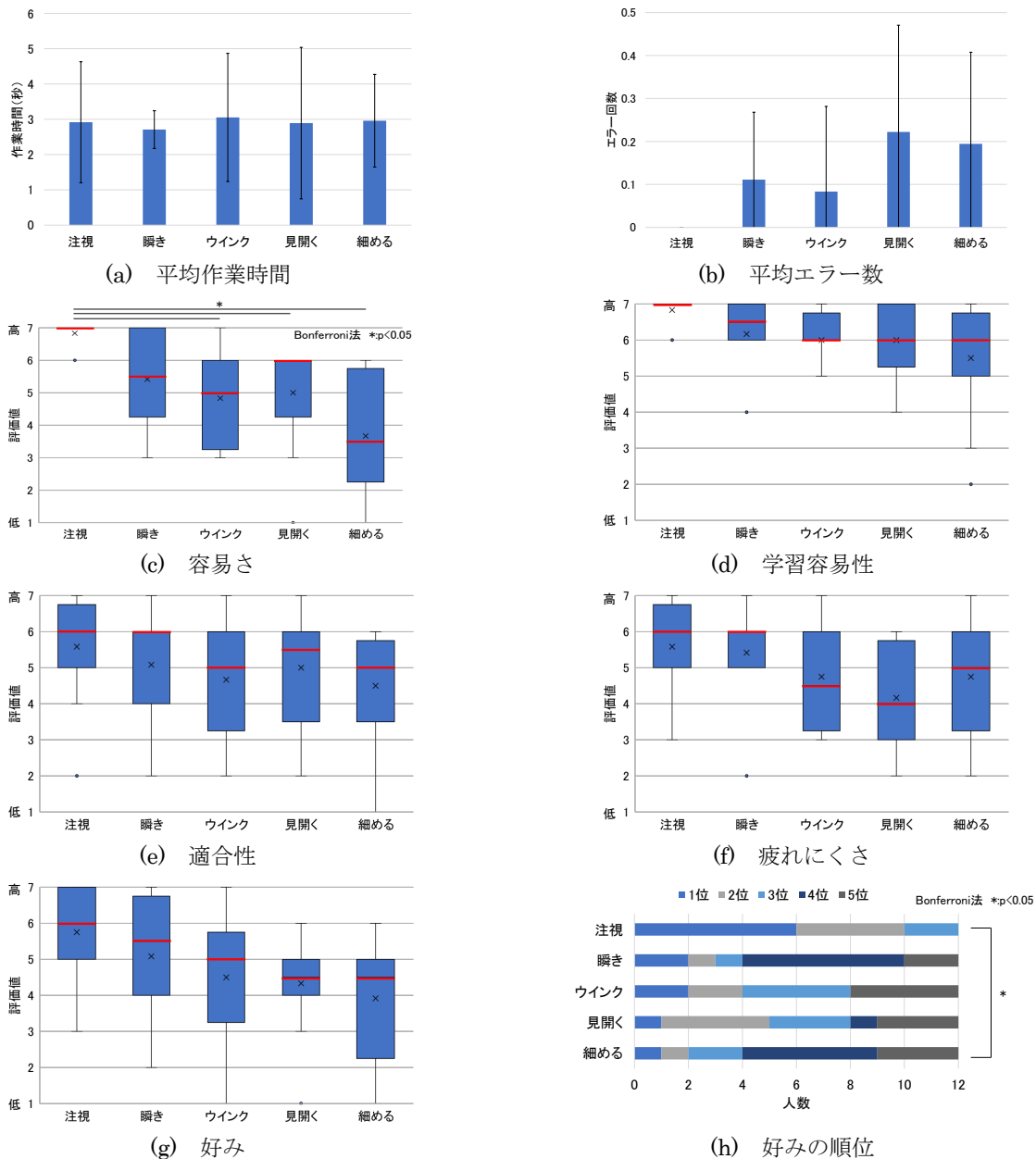


図7 タスク1：選択の結果

一ニ法を用いた多重比較を行った結果, 学習容易性(図8(d))ではウインクと細める, 適合性(図8(e))ではウインクと細める, 見開くと細める, 好み(図8(g))ではウインクと細める, 見開くと細めるの間に有意差が認められた ($p < 0.05$) . 入力方法の好みの順位(図8(h))に関しても, 入力方法間で有意差が認められた ($p < 0.05$) . ボンフェローニ法を用いた多重比較を行った結果, ウインクと見開く, 見開くと細めるの間に有意差が認められた ($p < 0.05$) .

4.5.3 タスク3：色変更

単発的入力&値変更操作では, 平均作業時間において, 入力方法の要因で主効果が有意であった ($F(4, 44) = 13.199, p < 0.05$) . 下位検定として, 入力方法間における有意差を確認するため, Ryanの方法による多重比較を行った. その結果, 注視と瞬き, 注視と細める, 瞬きとウインク, 瞬きと見開く, ウインクと細める, 見開くと細める

の間に有意差が見られた ($p < 0.05$) . 主観評価については, 容易さ(図9(c))では入力方法間で有意差が認められた ($p < 0.05$) . ボンフェローニ法を用いた多重比較を行った結果, 瞬きと見開くの間に有意差が認められた ($p < 0.05$) . 同様に, 適合性(図9(e))では注視と瞬き, 瞬きとウインク, 疲れにくさ(図9(f))では瞬きとウインクの間に有意差が認められた ($p < 0.05$) . 学習容易性(図9(d)), 好み(図9(g))に関しては, フリードマン検定において入力方法間における有意差は認められたものの, ボンフェローニ法を用いた多重比較を行った結果, 入力方法間において有意差は認められなかった. 入力方法の好みの順位では, 入力方法の要因における有意差は認められたものの ($p < 0.05$) , ボンフェローニ法を用いた多重比較を行った結果, 入力方法間において有意差は認められなかった.

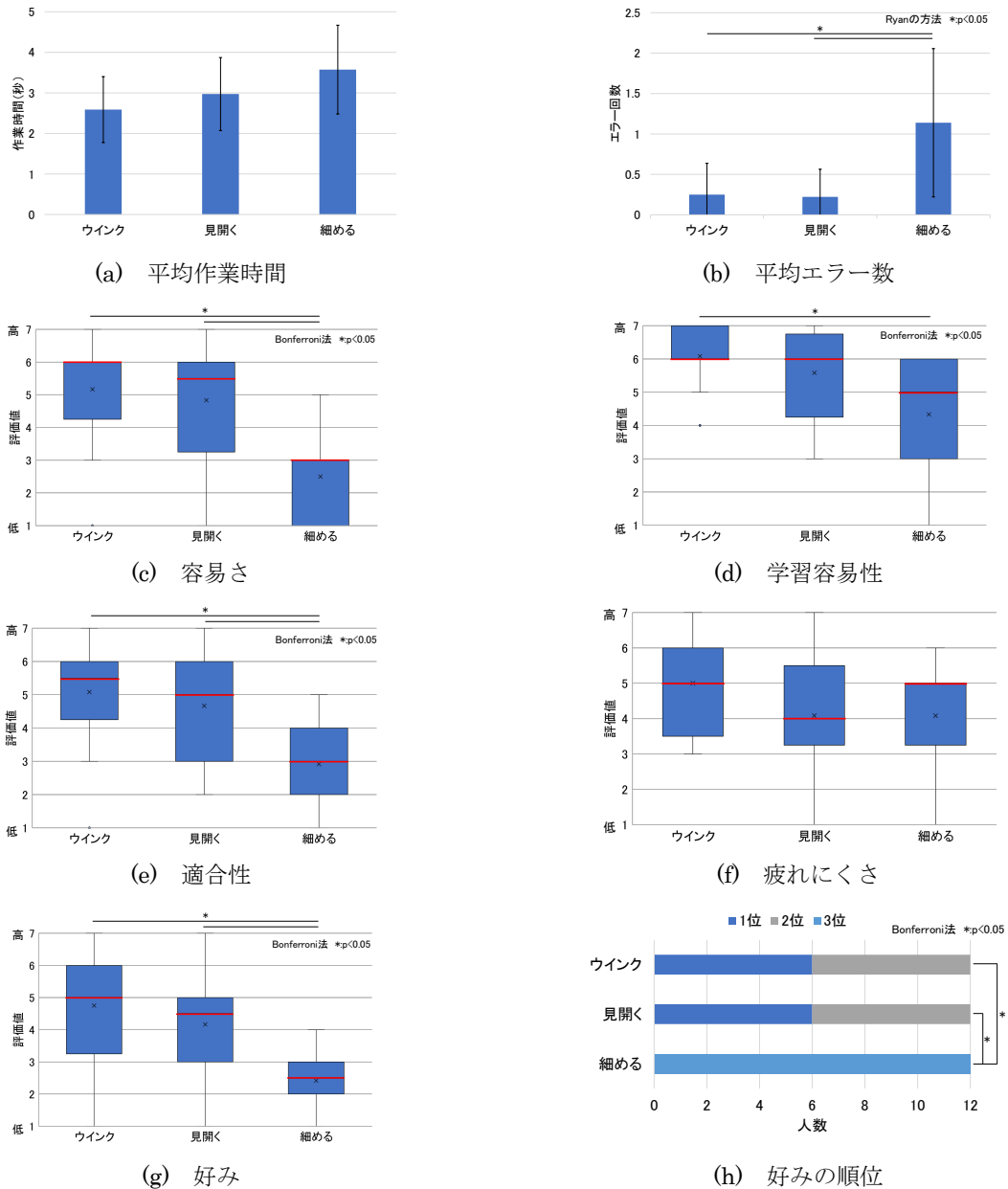


図8 タスク2: 掴みの結果

4.5.4 タスク4: 拡大

継続的入力&値変更操作では、主観評価、入力方法の好みの順位、いずれの結果についても、入力方法間で有意差は認められなかった。

5. 考察

4.5の実験結果から、VR空間操作コマンドとしての各アイジェスチャ(注視、瞬き、ウィンク、見開く、細める)にどのような特性があるのか、また、どのような操作に適しているのかについて考察する。

【注視】注視はどの実験においても主観評価項目の容易さと疲れにくさの中央値が5.5以上と高く評価されている。ただし、タスク1では他のアイジェスチャと比べて注視はどの項目でも高評価であったが、タスク3とタスク4では

評価が伸びなかった。これらのことから、注視は誰でも簡単に扱えて疲れにくい、複数回入力や継続的な入力には適していないことが示唆される。

【瞬き】瞬きはどの実験においても主観評価項目の容易さと疲れにくさの中央値が5.5以上と高く評価されている。特にタスク3の単発的な瞬きによる複数回入力(値変更操作)では他のアイジェスチャと比べてどの項目でも高評価であったが、タスク1の単発的な瞬きによる単数回入力(ONOFF操作)では評価が伸びなかった。これらのことから、瞬きは誰でも簡単に扱えて疲れにくい、単数回入力には適していないことが示唆される。

【ウィンク】ウィンクは今回評価した5つのアイジェスチャの中では最も苦手な人が多かった。タスク2では他のアイジェスチャと比べてウィンクは全体的に評価が高く最も好

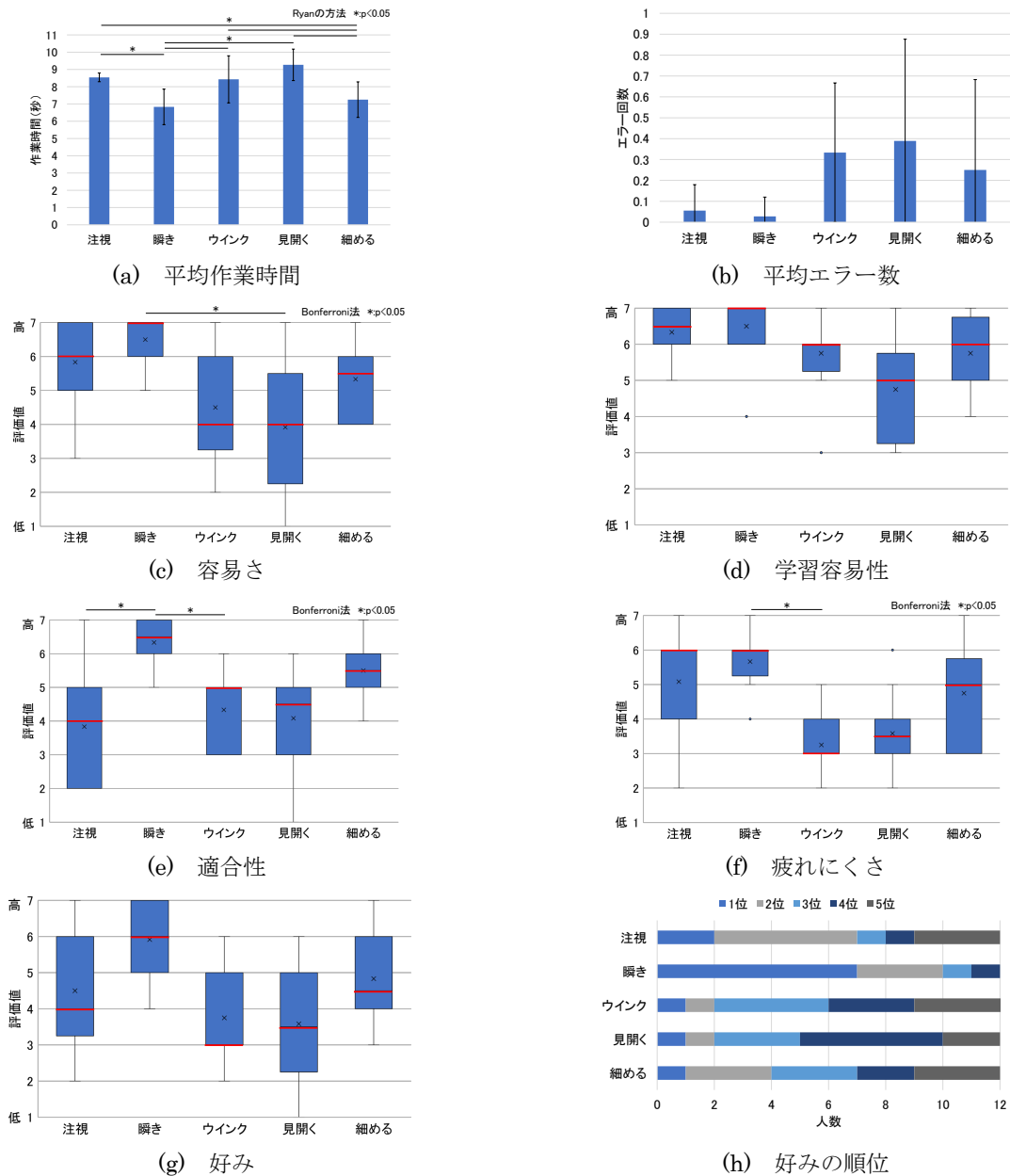


図9 タスク3:色変更の結果

まれていたが、タスク3では好みの評価が最も低かった。これらのことから、ウィンクは得意な人にはエラーが少なく使いやすいが、複数回入力に適していない、苦手な人には難しく疲れやすいことが示唆される。

【見開く】見開くは、継続的な入力を用いるタスク2, 4と比べて、単発的な入力を用いるタスク1, 3で好みの評価が低かった。また、見開くは全体に疲れにくさの評価が低かった。以上のことから、見開くは人によって得意不得意が分らない動作であるものの、単発的な入力には適していない、エラーが起りやすい、疲れやすいということが示唆される。

【細める】人によって得意不得意が分らない容易な動作であるがタスク1, 2のような視線を動かす必要のあるタスクにおいて、エラーが頻繁に発生した。これは、細めると

いうアイジェスチャは入力を行う際に目を細くする必要があり、これによってアイトラッキングデバイスが視線情報を正確に取得することができず視線がぶれやすくなるためである。

6. むすび

本研究では、各アイジェスチャに対して、VR空間における操作を4種類に分類したものをタスクとして設定した実験に取り組んだ。その結果を基に、各々のアイジェスチャの特性と、どのような操作に適しているのかを分析した。

まず、注視は単発的な入力を用いた単数回入力、瞬きは単発的な入力の値変更操作に適していた。ウィンクは得意な人には使いやすく、継続的な入力を用いたON/OFF操作に適していた。見開くと細めるに関しては、上記のアイジ

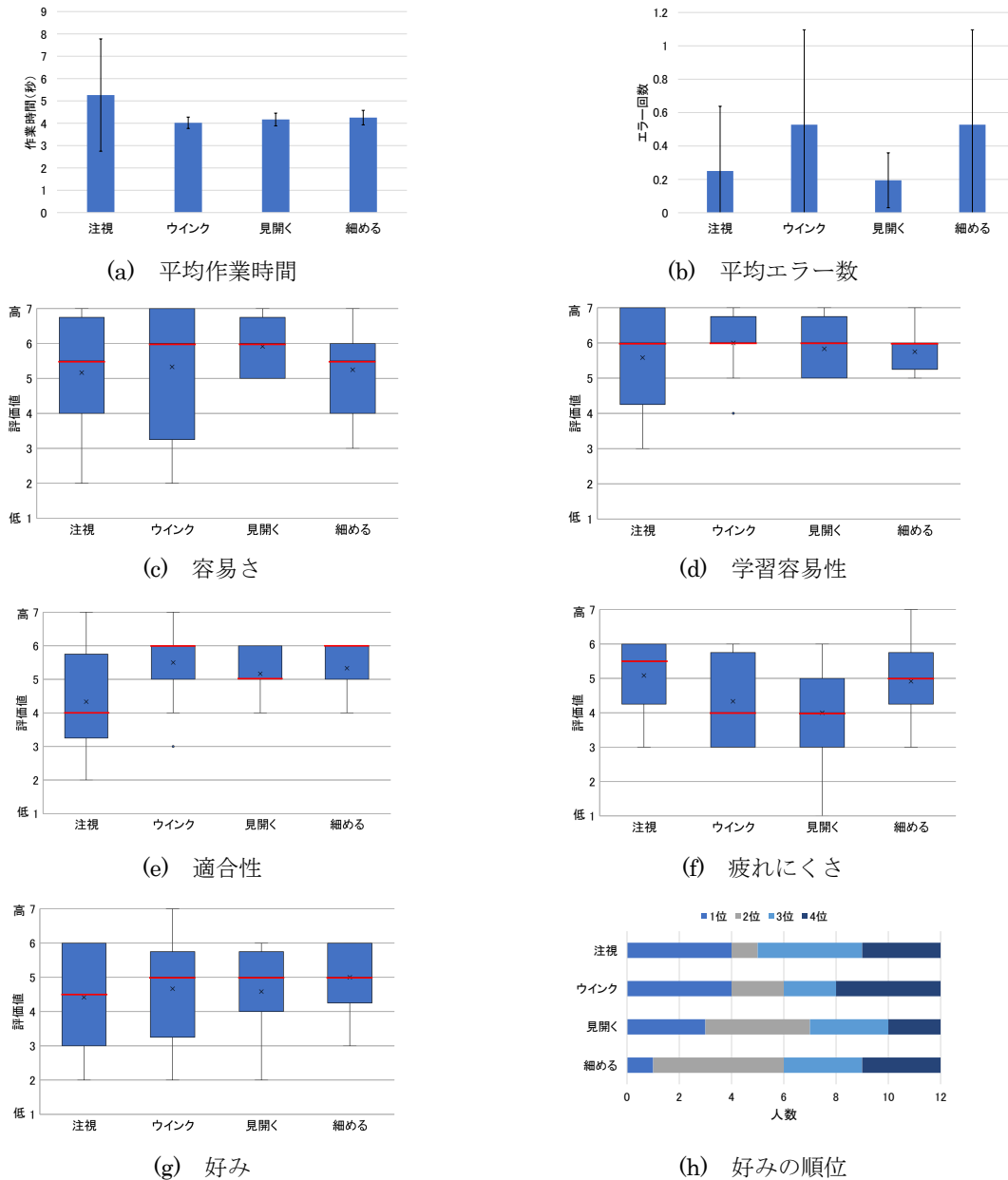


図 10 タスク 4 : 拡大の結果

エスチャよりも評価は下がるが、見開くは継続的な入力を用いた ON/OFF 操作に向いていると考えられ、細めるは強いて言えば継続的な入力を用いた値変更操作に向いていることがわかった。

最後に、本研究の使用機器を貸与していただいた京都大学工学研究科の中村裕一先生、ご指導いただいた井藤秀隆氏に感謝の意を表す。

参考文献

[1] J. S. Pierce, B. C. Stearns, and R. Pausch: "Voodoo dolls: seamless interaction at multiple scales in virtual environments," Proc. symposium on Interactive 3D graphics, pp.141 - 145, 1999.
 [2] F. Tecchia, G. Avveduto, R. Brondi, M. Carrozzino, and M. Bergamasco: "I'm in VR!: using your own hands in a fully immersive MR system," Proc. Virtual Reality Software and

Technology, pp. 73 - 76, 2014.
 [3] V. Rajanna, J. Hansen: "Gaze typing in virtual reality: impact of keyboard design, selection method, and motion," Proc. Eye Tracking Research & Applications, No. 15, 2018.
 [4] J. Orlosky, T. Toyama, K. Kiyokawa, and D. Sonntag: "ModulAR: Eye-controlled vision augmentations for head mounted displays," *IEEE transactions on visualization and computer graphics*, vol.21, No.11, pp. 1259 - 1268, 2015.
 [5] M. Choe, Y. Choi, J. Park and H. Kim: "Comparison of Gaze Cursor Input Methods for Virtual Reality Devices," *J. Human-Computer Interaction*, vol.35, pp. 620 - 629, 2019.
 [6] P. Ku, E. Bastias, T. Wu and M. Chen: "Wink It: Investigating Wink-based Interactions for Smartphones," Proc. Human-Computer Interaction with Mobile Devices and Services Adjunct, pp. 146 - 150, 2018.