

視線追跡 HMD と眉間を用いた「目掴み」の評価

池田沙厘奈[†] 薄羽大樹[†] 宮下芳明[†]

概要: 本稿では、視線追跡装置を搭載した HMD と眉間の動きを利用し、VR 空間内においてコントローラを用いずともドラッグ&ドロップが可能な「目掴み」を提案する。提案システムでは、ドラッグしたいオブジェクト上で眉間を寄せ、オブジェクトを選択し、そのままポインタを動かし（ドラッグ）、目的の領域上で眉間を離す（ドロップ）ことでドラッグ&ドロップが可能である。実験の結果、操作時間はターゲットまでの距離とターゲットの直径に影響され、エラー率はターゲットの直径に強く依存することが分かった。また、評価実験を踏まえ、システムの改善案を示した。

1. はじめに

近年、仮想現実感（Virtual Reality: VR）技術の普及と発展が急速に進んでいる。それに伴い、ヘッドマウントディスプレイ（Head Mount Display: HMD）も様々存在している。また、視線追跡の開発も進んでいる。これにより、ユーザがどこを注視しているのかを把握でき、注視情報を応用したインタラクション技術の開発が可能となってきた[1]。従来の研究では、PC のスクリーンを対象とした視線追跡が主であったが、今では、アイトラッカーが搭載された HMD も存在する。例えば、「StarVR one[2]」「Tobii Pro VR[3]」「FOVE 0[4]」である。HMD にアイトラッカーが搭載されることで、視野の中心付近だけを高解像度で見せる「フォービエイテッド・レンダリング」が可能になるなど、HMD においても視線を利用する技術が期待されている。

本稿第 3 著者らは、F.A.R Vision を提案した[5]。F.A.R Vision は、拡張現実感を対象としたシステムであり、シースルー型 HMD にフォトトリフレクタを取り付け、眉間の動きをセンシングしている。そのため、眉間の動きによってフェーダー操作を行うことが可能である。

これに対し、本稿では、VR 空間内において眉間の動きのよってドラッグ&ドロップが行える「目掴み」を提案する。提案システムでは、眉間を寄せることでオブジェクトを掴むことができ、眉間を寄せたままポインタを動かし、眉間を離すことでオブジェクトを離す。これによって、VR 空間内においても、コントローラを用いずともドラッグ&ドロップが可能である。また、素早く眉間を開閉すれば、クリック操作も行える。視線追跡が可能な HMD に実装されており、ユーザは視線と頭部の動きによってポインタを操作する。本稿では、ドラッグ&ドロップタスクによって「目掴み」の操作性についても調査する。

関連研究

1.1 視線追跡と HMD に関する研究

山浦らは、ユーザの視線に追従したエフェクトを重畳することでデジタルコンテンツの体験を拡張する手法を提案した[6]。エフェクトを重畳によって、そのコンテンツに対する印象の評価値が上昇することや、コンテンツにより集中できることが分かった。Thammathip らは、目の主要な 3 つの動きに注目し、視線追跡を用いた HMD の操作を提案した[7]。提案された操作は、(1) 対象物に視線を向ける高速で一過性の眼球運動であるサッケードを用いた Duo-Reticles, (2) スムーズな追跡を用いた Radial Pursuit, (3) 前庭動眼反射を用いた Nod and Roll である。

1.2 表情に関する研究

Chiba らは、従来の HMD がユーザの顔を隠してしまうことを問題視し、IR（赤外線）パスフィルタと IR カメラを用いて顔を撮影できる HMD を提案した[8]。鈴木らは、HMD の内側と外周部にフォトトリフレクタを組み込み、HMD を装着したユーザの表情を認識する技術を提案した[9]。提案手法によって、ユーザの表情をアバターへ反映できるようになった。(株) 360Channel は、FOVE 0 と口や顎の動きを取得できる Facial Tracking Camera によって表情を 3D アバターに反映させている[10]。Yamashita らは、光センサを HMD に取り付け、額[11]や頬[13]上でのタッチ操作を可能にした。菊池らは、光センサと機械学習を用いて、耳を入力インタフェースにするイヤホン型デバイス「Ear Touch」を提案した[12]。

1.3 ポインタに関する研究

山中らは、注視していない領域の特定が容易であることを応用し、注視していない領域ではポインタの速度を増加させる手法を提案した[14]。村田らは、視線入力装置でポインタを操作し、決定操作にはマウス、もしくは、視線入力（視線が一定時間停留した場合には決定操作になる）を用

[†] 明治大学

いたポインティング実験を実施した[15]. 実験によって、視線でのポインティングの精度を高めるには、20 pixels 以上のターゲット間隔、40 pixels 以上のターゲット幅にすべきであると示された。また、縦方向よりも横方向のポインティングの精度が高いことも示された。村田らは、同様の入力デバイスでドラッグ&ドロップの実験も実施した[16]. ドラッグ&ドロップでは、マウスを用いる方が視線入力よりも作業時間が短いことが示された。大和らは、視線の動きとマウスを併用した操作を提案した[17]. マウスのみの操作に比べて、操作の誤りを増やすことなく、平均で17% (0.4秒) 短く操作できることが示された。また初心者想定し、利き手ではない手を用いた実験も実施した。操作の誤りが減り、そして、平均で27% (0.8秒) 短く操作できた。曆本は、ドラッグ&ドロップ操作を複数のPCで可能にした、Pick and Dropを開発した[18]. PC上でPickしたデータをもう1台のPCにDropすることができ、ファイル転送コマンドを意識せずともファイルの移動が可能になった。

2. 提案システム

2.1 概要

ドラッグ&ドロップはGUI上で行われる頻繁な操作の一つである。例えばVR空間内では、まず、ドラッグしたいオブジェクト上でコントローラのボタンを押し、ボタンを押したままコントローラを動かす(ドラッグ)、目的の領域上でボタンを離す(ドロップ)ことで、ドラッグ&ドロップが行える。これに対し、提案システムでは、眉間を寄せているか・離しているかをセンシングし、その情報を用いることでドラッグ&ドロップを実現する。つまり、提案システムでは、VR空間内において、コントローラを用いずともドラッグ&ドロップが行える。

提案システムでは、まず、ドラッグしたいオブジェクト上で眉間を寄せ、オブジェクトを選択し、そのままポインタを動かす(ドラッグ)、目的の領域上で眉間を離す(ドロップ)ことで、ドラッグ&ドロップが行える。つまり、しながら、眉間でオブジェクトをつかんでいるような操作が行える。マウス操作で例えるなら、眉間を寄せることがマウスダウンであり、眉間を離すことがマウスアップにあたる。また、ポインタの操作は頭部の動きと視線によって行う。

提案システムの応用例として、クリック操作があげられる。近年では、「OptiKey[19]」などの視線計測を用いた文字入力インタフェースが存在する。OptiKeyでは、視線でポインタを動かす、マウスやコントローラなどのボタンを押すことで文字の選択が可能である。一方で、提案システムでは、視線や頭部の動きによってポインタを動かす、素早く眉間を開閉することでクリック操作が行え、文字の入力が行えると考えている。

ポインタは注視点に黒色で表示されている。眉間が寄っている状態では、ポインタの色は黒色から緑色に変わる。

2.2 構成

提案システムの構成を図1に示す。FOVE 0の眉間にあたる位置に穴を空けてフォトリフレクタを内蔵した(図2)。これにより、眉間を寄せているか・離しているかをセンシングすることが可能である。また、FOVE 0からは視線と頭部の動きを取得する。そして、Unityでそれらの情報を統合し、ポインタなどをFOVE 0内のスクリーンに描画する。FOVE 0の解像度は2560×1440 pixels、フレームレートは70 fps、視野角は100度、視線追跡器のトラッキング精度は1度未満、フレームレートは120 fpsである。

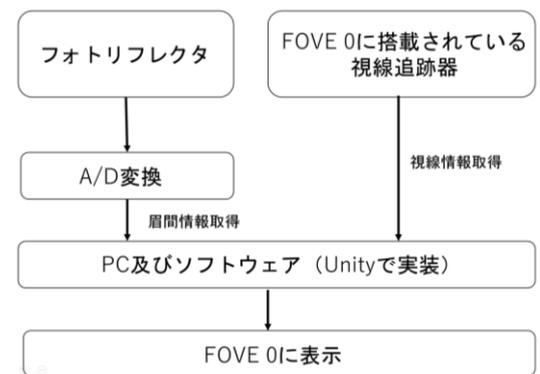


図1 提案システムの構成



図2 FOVE 0と取り付けられたフォトリフレクタ

視線のキャリブレーションについては、FOVE 0のシステムを用いた。また、眉間のキャリブレーションについては、F.A.R Vision[5]での手順を用い、2セット行った。F.A.R Visionのキャリブレーション方法とは、まず、計測した2セットのデータから各セットの最大値と最小値を求める。次に得られた2セットの各最大値と最小値から「2セットの中で小さい最大値」と「2セットの中で大きい最小値」を求め、その値を出力値の最大と最小として認めた。そして検出幅はその差と定義した。3つのフォトリフレクタのうち、検出幅が一番大きいものを採用した。閾値は、検出幅の半分の値を最小値に足したものとした。

3. 実験

3.1 参加者

参加者は、21~24歳の男子大学生7名であった。参加者のうち2名は眼鏡を日常的に使用しているが、眼鏡を装着した状態でFOVE 0のキャリブレーションを行うことがで

きなかったため、裸眼で実験を行った。

3.2 タスク

参加者が行うタスクの内容を図 3 に示す。まず、参加者は開始ターゲット（赤色）上まで視線と頭部によってポインタを動かし、眉間を寄せ、試行を開始させた。その後、眉間を寄せたまま、視線と頭部でポインタを動かし、終了ターゲット（黄色）を目指し、眉間を離すことでその試行が終了する。眉間を離れたとき、ポインタの座標が終了ターゲット上であった場合には、その試行を成功とした。試行が成功した場合にはベル音が再生された。また、それ以外の場合（例えば、途中で眉間が離れた場合、終了ターゲット外で眉間が離れた場合）はエラーとし、失敗を伝える音が再生された。

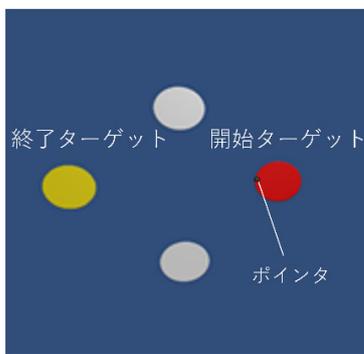


図 3 実験システムの画面

3.3 デザイン

実験デザインの概要を図 4 に示す。開始・終了ターゲットの直径 (W) は、1.2 m (視角 9.15°)、1.5 m (視角 11.42°)、1.8 m (視角 13.69°)、開始ターゲットの中心から終了ターゲットの中心までの距離 (D) は 4.78 m (視角 35.35°)、5.75 m (視角 41.95°) であった。開始ターゲットから終了ターゲットまでの方向は 4 種類であり、上から下、下から上、右から左、左から右であった。VR 空間内では、カメラからターゲット面までの距離は 7.5m であった。1 セットは $W(3) \times D(2) \times \text{方向}(4) = 24$ 試行であり、 W 、 D 、方向の出現順序はランダムであった。

3.4 手順

まず、参加者に実験の内容を説明し、眉間と視線のキャリブレーションを行った。そして、参加者は 1 セットの練習を行い、その後、10 セットを実験用のデータとして行った。つまり、全体では $W(3) \times D(2) \times \text{方向}(4) \times 10 \text{ セット} \times 7 \text{ 名} = 1680$ 回のデータが集計された。実験終了後、提案システムに関してのアンケートを実施した。それぞれの参加者が実験に要した時間は 30~50 分であった。参加者は回転可能な椅子に着席しており、ターゲットを選択する際には視線を動かしても頭を動かしても良いと伝えた。また、実験中、参加者が眼疲労や酔いを訴えた場合は休憩や中断が可能であることを伝えた。

3.5 計測項目

開始ターゲット上で眉間を寄せてから、終了ターゲット上で眉間を離すまでの時間 (MT)、エラー率を記録した。

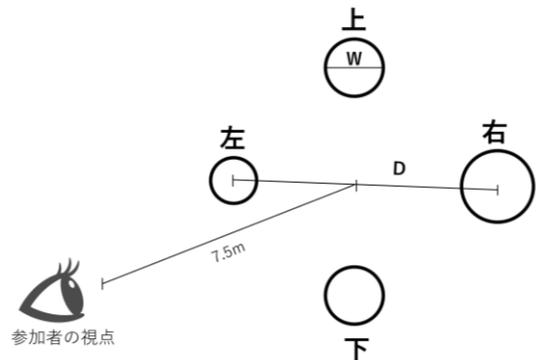


図 4 実験デザイン

4. 結果

1680 回のデータのうち、6 回のデータが欠損した。また、エラーした試行のうち、ポインタが開始ターゲット付近で終了している 81 個のデータは、開始直後に眉間を離してしまった試行であるため、外れ値として除いた。最終的には、実験データとして 1593 回のデータを用いた。独立変数を W 、 D 、方向、従属変数を MT とエラー率とし、分散分析を行った。その後の検定には Bonferroni 法を用いた。全体のエラーは 358 回 (22.5%) であった。グラフ中のエラーバーは標準誤差を示す。多重比較によって有意差が見られた場合には、グラフ中に p 値を記述している。

4.1 D と W と方向の MT への影響

操作時間に関する結果についてはエラー試行を排除した上で計算を行った。 D ($F_{1,6} = 44.13, p < 0.01$) と W ($F_{2,12} = 4.58, p < 0.05$) では、 MT に主効果があることが示された。一方、方向 ($F_{3,18} = 0.47, p > 0.10$) には、主効果が見られなかった。また、多重比較の結果、 W が小さくなるほど ($W = 1.2 \text{ m}$ と $W = 1.8 \text{ m}$ 間では $p < 0.05$) MT が増加した (図 5)。

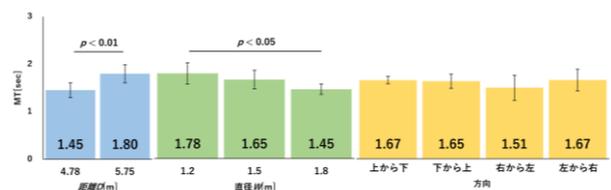


図 5 D と W と方向の MT への影響

4.2 エラー率

W ($F_{2,12} = 9.61, p < 0.01$) では、エラー率に主効果があることが示された。一方、 D ($F_{1,6} = 0.28, p > 0.10$) と方向 (F

3,18 = 0.14, $p > 0.10$) には主効果が見られなかった。また、多重比較の結果、直径 W が小さくなるほど ($W=1.2\text{ m}$ と $W=1.8\text{ m}$ 間, $W=1.5\text{ m}$ と $W=1.8\text{ m}$ 間では $p < 0.05$, それ以外は $p > 0.10$) エラー率が増加した (図 6)。

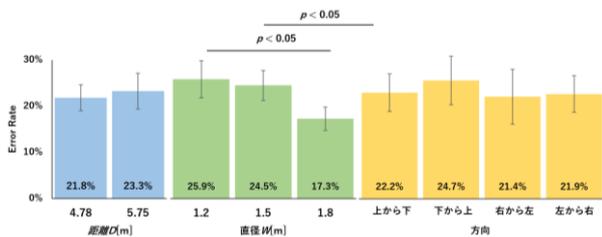


図 6 D と W と方向のエラー率への影響

また、図 7 に示すように、 $D \times$ 方向では交互作用が見られ ($F_{3,8} = 3.80, p < 0.05$), 左から右に操作を行う場合には、 D が小さいほどエラー率が減少することが観察された。

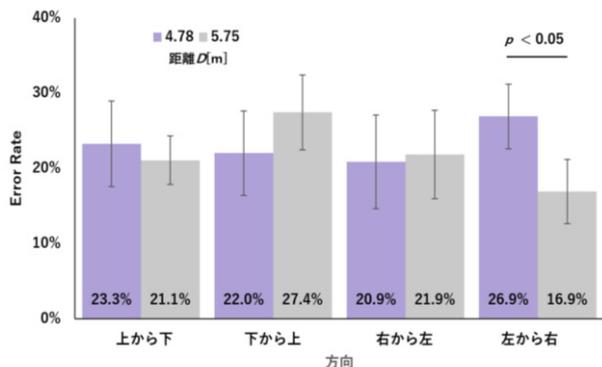


図 7 D と各方向のエラー率

4.3 アンケート

実験終了後、アンケートを行った。日常的に HMD で VR 体験をしているか、提案システムの操作はしやすかったか、また、システム使用後の目の疲労と酔い、操作がしやすかった方向・しづらかった方向、提案システムについての自由記述を回答してもらった。7名中1名が、普段から VR 体験をしていると答えた。操作のしやすさ、目の疲労、酔いについては図 8 に示す。

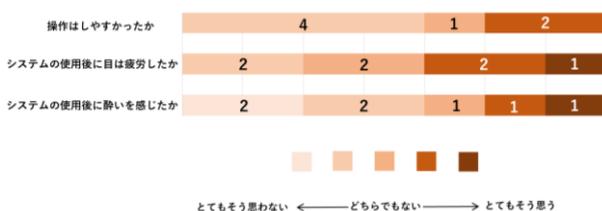


図 8 アンケートの結果

視線と眉間によるドラッグ&ドロップについては、操作しづらいと答えた人が多かった。目の疲労については、参

加者によって回答は異なり、個人差があると考えられる。また、実験中 30~50 分 HMD を装着していたのにも関わらず、酔いにくいことが分かった。操作がしやすかった方向・しづらかった方向については、7名中4名が「眉間を寄せたままだと上下の動きがしづらい」と答えた。「上下よりも左右の動きがやりやすかった」、「上を見たときに眉間の寄せが外れてしまったかもしれない」という回答も得られ、主観的には、上下方向のドラッグ&ドロップが難しいことが分かった。自由記述欄では、「視線計測がもっと正確であれば十分に使えると感じた」、「眉間を動かすコツを途中でつかんでから操作が楽になった」という意見があった。また、実験では眉間を使って 240 回の試行をするため「眉間で長時間操作するのは、普段あまり使わない筋肉を使うので疲れた」という意見もあった。

5. 考察

5.1 視線のぶれと眉間キャリブレーション

MT が大きくなる・エラー率が高くなる原因は、ターゲットの直径などの実験条件以外では、眉間を寄せることによる視線のぶれであると考えられる。FOVE0 の視線計測は、目に反射する赤外線を用いて行っている。そのため、眉間を寄せ、目が細くなると赤外線を反射できなくなるため、視線計測が難しくなり、計測される視線にぶれが生じる。ターゲットの直径が小さい場合には、ぶれを考慮しながらポインタをターゲット上にのせる必要があるため、ドラッグ&ドロップがより難しくなる。本実験では、視線と頭部の動きによってポインタ操作してもらったが、ポインタを常にスクリーンの中央に配置し、頭部の動きのみで操作する方法も考えられる。また、赤外線を用いた視線計測ではなく、眉間周辺の映像をカメラで取得し、画像処理によって視線を推定できれば、たとえ眉間を寄せたとしても視線がぶれることなく操作が行えるだろう。

提案システムでは、HMD を一度外し再び装着すると、HMD 内に入る光の量に変化し、また、フォトリフレクタと眉間の位置もずれてしまう。そのため、フォトリフレクタから得られる値が大幅に変わってしまい、キャリブレーションを再度行う必要があった。また、眉間にしわが寄りやすい参加者・寄りにくい参加者がいた。これらのことを考慮すると、眉間のセンシングに関しては、より精度を上げる必要があると考えられる。

5.2 D と W と方向の影響

ターゲットまでの距離 (D) とターゲットの直径 (W) が眉間によるドラッグ&ドロップの操作時間 (MT) へ影響を与えることが分かった。 D に関しては、ターゲット間の距離が小さくなるほど MT が小さくなることが分かった。 W に関しては、1.2 m と 1.5 m 間、1.5 m と 1.8 m 間には差がなかった。5.1 節でも述べたが、眉間を寄せた状態の視線計測は難しく、ポインタがぶれ、それが影響したと考えられ

る。そのため、 $W=1.2\text{ m}$ と $W=1.8\text{ m}$ の間では大きく MT に差がついたと考えられる。一方で、アンケートでは「上下の動きがしづらい」という回答があったものの、方向に関しては主効果が見られなかった。

5.3 エラー率

図 6 に示すように、 W がエラー率へ影響を与えることが分かった。5.1 節で述べたように、眉間を寄せたままであるとポインタが安定しないということが、エラー率にも影響したと考えられる。 $W=1.2\text{ m}$ と $W=1.8\text{ m}$ 間、 $W=1.5\text{ m}$ と $W=1.8\text{ m}$ 間では差があったことから、眉間を用いた手法においては、ターゲットの直径が $W=1.8\text{ m}$ 以上あるとエラー率が少なく、安定して操作できると考えられる。一方で、Kumar ら[21]によれば、PC 画面上で視線ポインティングを行う場合のエラー率は、ターゲット幅が 22 pixels のとき 16%、33 pixels のとき 17%、44 pixels のとき 15%であることが分かっている。我々と Kumar らの実験結果を比較すると、 $W=1.8\text{ m}$ 以下のターゲットにおいても、より低いエラー率で操作できる方が理想であると考えられる。

$D \times$ 方向においては、相互作用が見られた。一方で、 $MT \times$ 方向では相互作用は見られなかったが、各方向での D ごとのエラー率を見てみると、 $D=4.78\text{ m}$ のとき $MT=1.35\text{ sec}$ 、 $D=5.75\text{ m}$ のとき $MT=2.00\text{ sec}$ であった(図 9)。差はなかったものの、他方向に比べ $D=4.78\text{ m}$ のときの MT は小さく、 $D=5.75\text{ m}$ のとき MT は大きくなっていった。つまり、 $D=4.78\text{ m}$ のとき、 MT が小さいためエラー率が高くなり、 $D=5.75\text{ m}$ のとき、 MT が大きいためエラー率が低くなっていたと考えられる。実験条件に斜めの方向を加え、4 方向から 8 方向に増やした実験を行うことで、ドラッグ & ドロップの性能と方向の関係がより明らかになるだろう。

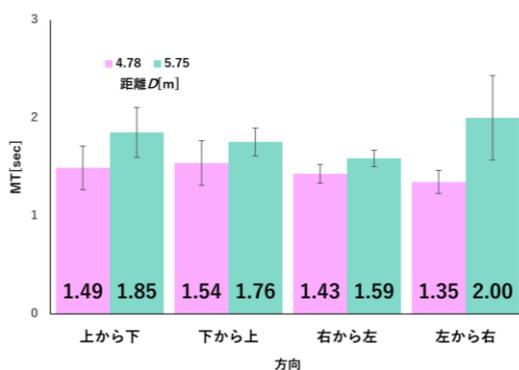


図 9 D と方向の MT への影響

6. 結論・課題

本論文では、視線追跡装置を搭載した HMD FOVE 0 に眉間をセンシングする機構を加えた「目掴み」を提案し、その操作性についての調査を行った。眉間を寄せることでオブジェクトを選択し、眉間を寄せたまま視線と頭部の動き

によってポインタを動かし、目的の領域で眉間を離すことでドラッグ&ドロップ操作が行える手法である。

実験により、ターゲットまでの距離 (D) が小さくなるほど、また、ターゲットの直径 (W) が大きくなるほど操作時間 (MT) が減少した。そして、エラー率に関しては、 D や向きの影響は少なく、 W に強く依存することが分かった。また、眉間を寄せることによる視線のぶれが MT やエラー率に影響していることも明らかになった。

今後、視線のキャリブレーションについては、赤外線光によるものではなく画像処理で行う方法など検討し、視線のぶれを改善する必要がある。また、より様々な方向などを加えた実験によって、眉間によるドラッグ&ドロップの性能を明らかにしていく予定である。

参考文献

- [1] R.J.K. Jacob. What you look at is what you get: eye movement-based interaction techniques. CHI 1990, pp.11-18, 1990.
- [2] “StarVR one”. <https://www.starvr.com/products/> (参照 2018-10-19) .
- [3] “Tobii Pro VR Integration”. <https://www.tobii.com/ja/product-listing/vr-integration/> (参照 2018-10-19) .
- [4] “FOVE Rye Tracking Virtual Reality Headset: Home”. <https://www.getfove.com/> (参照 2018-10-19) .
- [5] Hiromi Nakamura, Homei Miyashita. Control of augmented reality information volume by glabellar fader. AH2010. Vol.20, pp.1-3, 2010.
- [6] 山浦祐明, 田村征優紀, 中村聡史. 視線に追従したエフェクト重畳によるデジタルコンテンツの体験拡張手法の提案. 研究報告ヒューマンコンピュータインタラクション, Vol. 2016-HCI-169, No. 13, pp.1-8, 2016.
- [7] Thammathip Piumsomboon, Gun Lee, Robert W. Lindeman, Mark Billingham. Exploring natural eye-gaze-based interaction for immersive virtual reality. 3DUI2017, 2017.
- [8] Mariko Chiba, Wataru Yamada, Hiroyuki Manabe. Transparent Mask: Face-Capturing Head-Mounted Display with IR Pass Filters. UIST2018, pp.149-151, 2018.
- [9] 鈴木克洋, 中村文彦, 正井克俊, 伊藤勇太, 杉浦裕太, 杉本麻樹. AffectiveHMD: 組み込み型反射型光センサを利用した表情認識とバーチャルアバターへの表情のマッピング. バーチャルリアリティ学会論文誌, Vol.22, No.3, pp. 379-389, 2017.
- [10] “FACE ~Face And Communication Entertainment~”. <https://face.360ch.tv/> (参照 2018-10-19) .
- [11] Koki Yamashita, Yuta Sugiura, Takashi Kikuchi, Maki Sugimoto. DecoTouch: Turning the Forehead as Input Surface for Head Mounted Display. ICEC2017, pp.481-484, 2017.
- [12] 菊地高史, 正井克俊, 杉本麻樹, Bruce H. Thomas, 杉浦裕太. EarTouch: 耳の入力インタフェース化. 日本バーチャルリアリティ学会論文誌, Vol.22, No.4, pp.477-483, 2017.
- [13] Koki Yamashita, Takashi Kikuchi, Katsutoshi Masai, Maki Sugimoto, Bruce H. Thomas, Yuta Sugiura. CheeckInput: turning your cheek into an input surface by embedded optical sensors on a head-mounted display. VRST2017, No.19, 2017.
- [14] 山中 祥太, 栗原 一貴, 宮下 芳明. 注視していないことを

利用したマウスカーソル高速化手法. WISS2012, pp.127-132, 2012.

- [15] 村田 厚生, 三宅 貴士, 森若 誠. 視線入力システムによるメニュー選択方法の有効性－若年者と高齢者の比較－. 人間工学, Vol.47, No.1, pp.20-30, 2011.
- [16] 村田厚生, 三宅貴士, 森若誠. 視線入力システムによるポイント操作時のクリック方法とドラッグ&ドロップ方法の提案－若年者と高齢者の比較－. 人間工学, Vol.47, No.1, pp.10-19, 2011.
- [17] 大和正武, 神代知範, 門田暁人, 松本健一. 視線・マウス併用型インタフェースのドラッグ&ドロップ操作への適用. 情報処理学会論文誌, Vol.44, No.1, pp.166-175, 2003.
- [18] 暦本純一. Pick-and-Drop: 複数コンピュータ環境での直接操作技法. WISS1997, pp.141-150, 1997.
- [19] “OptiKey” . <https://github.com/OptiKey/OptiKey/wiki> (参照 2018-12-19) .
- [20] 神代知範, 大和正武, 門田暁人, 松本健一, 井上克郎. ドラッグ&ドロップの操作時間を視線情報とマウスを併用して短縮する方法. 第 59 回全国大会講演論文集 (インタフェース) , pp.33-34, 1999.
- [21] Manu Kumar, Andreas Paepcke, Terry Winograd. EyePoint: Practical Pointing and Selection Using Gaze and Keyboard. CHI2007, pp.421-430, 2007.