
視線とハンドジェスチャーを併用したポインティング機能の実装

システム理工学部教授 榎原博之
理工学研究科 直野智仁

1 はじめに

近年、センシング機器の多機能化、高性能化、低価格化により、ジェスチャーや音声などを利用した様々な入力機器が販売されている。従来の入力方式であるマウスなどの入力方式は接触型インターフェースと呼ばれる。接触型インターフェースはデバイスとの物理的接触を必要とするため、衛生面などから使用できない問題がある。一方現在、非接触型と呼ばれる物理的接触を必要としない、新たな入力方式が研究されている。

非接触型インターフェースの一つに、アイトラッキング技術を用いて視線で入力を行うものが存在する。アイトラッキングとはユーザがどこを見ているかを測定する技術である。この技術は、マーケティング調査やゲーム、心理学や医学などの多岐の分野に渡って活用されており、活用分野の一つには、アイトラッキングを利用した GUI 操作の研究がある。視線入力は、非常に速い眼球運動で、最大速度は1000 deg/s になることもある。また、人間がクリックなどの操作を行う際に選択対象を見てからカーソル移動を行うという特性上、注視点をカーソル座標とすることで高速なポインティングが行えることがわかっている。しかし、視線入力は小さな選択対象にはずれが生じるため微調整が効かない、クリックのような選択操作には凝視やまばたきが一般的だがユーザへの負担が大きい、視線は不随意のノイズが混ざりやすい、といった欠点がある。

また、視線以外の非接触インターフェースには、ジェスチャー入力が存在する。ジェスチャー入力は身振り手振りによって入力を行うことができるため、直感的な入力方式として注目されている。しかし、ジェスチャー入力の特性上、ジェスチャーの種類が増えるほど誤識別が増える、操作回数が増えるほどユーザの疲労が大きくなる、といった欠点がある。

そこで、本提案システムでは選択対象までの大まかな移動は視線で行い、最後の微調整と選択操作をジェスチャーで行うことで、視線の俊敏性を活かしながら小さな対象の選択も行うことができるシステムを提案し、実装する。

2 非接触型インターフェース

2.1 アイトラッキング

アイトラッキングとは、使用者がどこをどのように見ているのかを測定する技術である。マーケティング調査やゲーム、ヒューマンインターフェース、心理学や医学など様々な分野において活用されている。

パソコンを操作する際に、ディスプレイ上における使用者の視線情報をリアルタイムに取得することで、視線の移動に伴うカーソル移動や、注視や瞬きによるターゲット選択など、視線情報をインターフェースに利用することができる。

眼球運動には、固視微動¹⁾やサッカード²⁾といったものがある。固視微動は、一点を注視しているとき、眼球が静止しているように見えても、常に細かく動いている眼球運動のことである。固視微動は単なるノイズではなく、眼球の方向を一定に保つための動きや、網膜像を鮮明に保つための動きと考えられている。サッカードは、本を読んでいる次の行に移るときなど、注視対象を変えるときに生じる。これは、非常に速い眼球運動で、最大速度は1000 deg/s になることもある。

GUI 操作において、ターゲット選択は頻出する作業である。ターゲットのポインティングを行う際、そのターゲットの位置を注視してからカーソルを動かすことが多いことが知られている³⁾。サッカードの速度の方が手でマウスを扱うよりも明らかに速いので、視線でポインティングを行うとポインティング速度が向上する。しかし、現状の OS やブラウザにおける GUI 上のターゲットサイズは小さく、固視微動の影響によってカーソルがぶれるため、選択の決定を行うのは難しい。

選択を行う操作方式としては、ターゲットの注視や瞬きなどが存在する。注視での決定方式では、ある一定時間の間、ターゲットを注視し続けることで選択が行われる。選択が決定されるまでの時間が短いと、操作時間は速くなるが、視線を止めるとすぐに選択が決定されてしまうため、意図しない選択ミスが生じてしまう。決定されるまでの時間を長くすると、選択ミスは減少するが、視線を一点に集中させておく時間が長くなり、目が疲れる。この他にも、選択決定のための注視であるのか、単に画面を眺めているだけなのか判別できないという問題点がある。また、一定時間の間、視線がターゲット上に居続けないといけませんが、固視微動の影響によりぶれてしまう。そのため、ぶれてもターゲット上に収まるようにターゲットサイズを大きくするなどの工夫が必要になる。瞬きによる決定操作は、ターゲット上に視線を合わせた上で瞬きをすると選択の決定が行われる。ターゲットを見続ける必要がないため、注視よりも操作時間が短く、目が疲れにくい。しかし、意図的な瞬きと生理的に起こる瞬きの判別が難しいため、選択ミスが生じる。

2.2 ジェスチャーインターフェース

本節では、ジェスチャー取得機器と関連研究について説明を行う。ジェスチャーインターフェースとは、パソコンやロボットなどを身振り手振りで操作するユーザインターフェースである。ジェスチャーインターフェースには、非接触でジェスチャーを取得するものがある。この手法は、画像認識やサーモグラフィーを用いて体の状態を取得するため、機器との物理的接触を必要としない。非接触型のジェスチャーの利点として、高精度で検出が可能であるということが挙げられる。例えば、画像認識を用いた機器には、Leap Motion Controller (LMC)⁴⁾という赤外カメラを用いた機器が一般社会でも普及している。LMCはハンドジェスチャーに特化した機器であり、指を関節ごとに識別可能な程精度が高く、ジェスチャーの設定を細かく行うことができる。また、非接触であるため無菌性が高く、医療など衛生環境に対して厳しい条件のある環境でも使用できる。一方で、短所としてカメラに対するユーザの距離が適切でない場合にはジェスチャーの取得精度が大きく低下する問題もある。例えば、LMCではカメラに手が近すぎる場合、手の一部だけしかカメラに映らず、手だと認識できない。

3 提案システム

本節では提案システムの実装と評価実験について説明する。本提案システムでは、大まかなポインティングを視線で行い、細かい微調整やクリックなどの選択操作をハンドジェスチャーで行うシステムである。前提として、本提案システムはマウスポインティングに代わるものではなく、料理中などの衛生的な観点から、入力装置への物理的接触を制限された状況下での使用を想定している。

提案システムの流れを図1に示す。提案システムは、大まかに分けると視線取得部、ジェスチャー取得部の二部で構成されている。図1に示すように、提案システムではまず手の情報を取得し、設定したジェスチャーかどうかを判別する。ジェスチャーの検出後、設定したジェスチャーが発生していなかった場合、視線情報の取得を行う。機器から視線情報が取得できなかった場合には、ジェスチャーの取得に戻る。視線情報が取得できた場合、取得した視線情報には固視微動というノイズが混ざっている。そのため、1点を見ているつもりでもカーソルは絶えず振動し、選択の精度が低下する。この固視微動を抑えるため、ローパスフィルタの一種である指数移動平均フィルタを用いてノイズを軽減する。その後、フィルタリングした視線座標を次のカーソル座標として更新を行う。

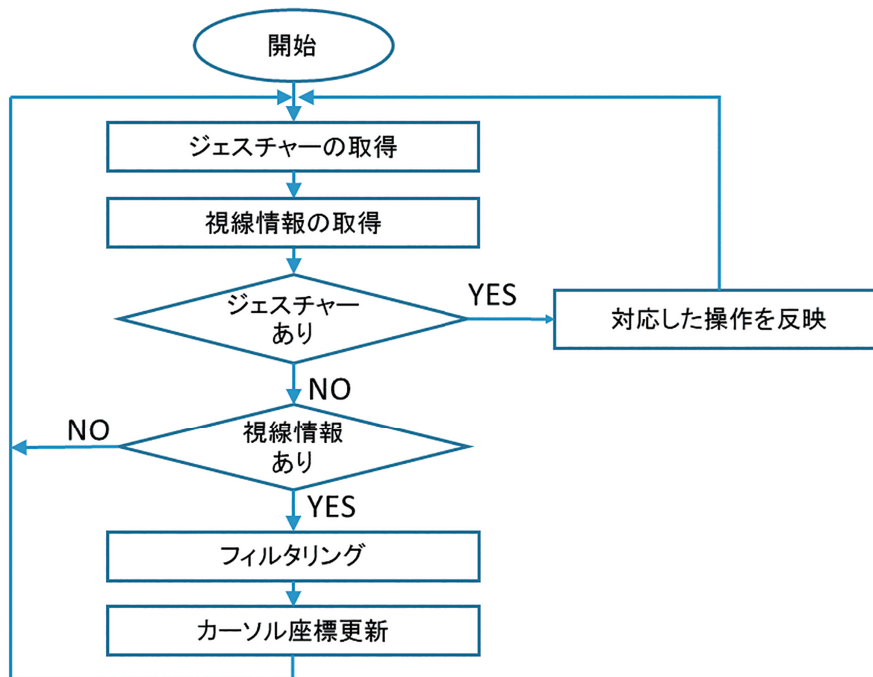


図1 提案システムの流れ

3.1 ジェスチャー取得部

今回、ジェスチャーの取得には LMC (Leap Motion Controller) を用いた。Leap Motion Controller は2012年に Leap Motion 社から発売された手のジェスチャーによる PC 操作を可能にする入力デバイスであり、マウスや、タッチパネルを使用せずに体感的、直感的に操作することができる。LMC では、指、手のひら、手首、肘のそれぞれの情報が取得可能である。また、指の取得情報には図2に示すように、5本の指すべてについて、第1関節から第4関節までの情報が取得可能なため、任意のハンドジェスチャーを細かく設定することができる。LMC で取得できる情報は相対位置、方向、計測時刻の3種類である。相対位置と方向

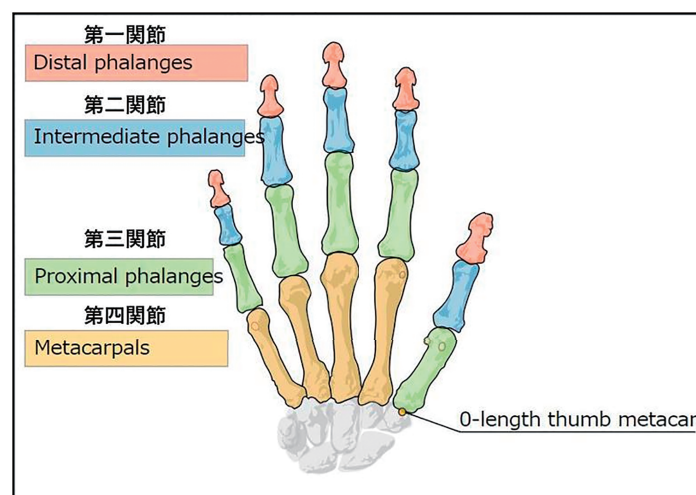


図2 LMC の指関節

(出典 : https://developer-archive.leapmotion.com/documentation/java/devguide/Leap_Overview.html)

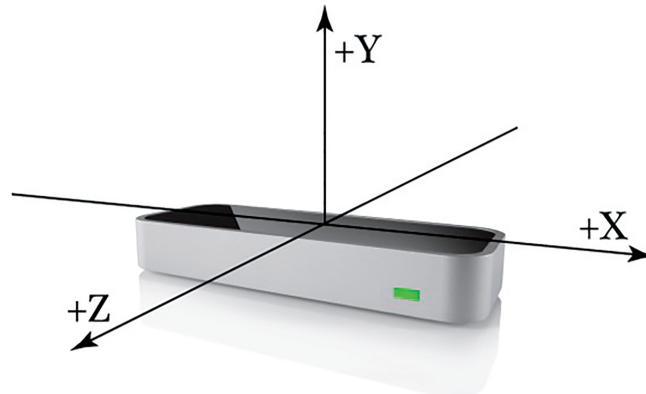


図3 LMCの座標系

(出典：https://developer-archive.leapmotion.com/documentation/java/devguide/Leap_Overview.html)

は図3に示すようにカメラを中心とした3次元空間上の座標、方向として出力される。このLMCを用いてクリック、ドラッグ&ドロップ、カーソル移動の3種類のジェスチャーを設定する。3種類のジェスチャーは、以下のようにジェスチャーの検出を行っている。

1. クリック

タップ動作でクリックを行う。人さし指の第二関節の方向を取得し、z軸のパラメータが閾値を下回った場合にクリックを行う。

2. ドラッグ

握り拳を作ることでドラッグする。中指の第二関節の方向を取得し、軸のパラメータが閾値を下回った場合にドラッグを行う。

3. ドロップ

握り拳の状態から、手を開くことでドロップする。中指の第二関節の方向を取得し、軸のパラメータが閾値を上回った場合にドロップを行う。

4. カーソル移動

LMCと手のひらの相対位置によって、カーソル移動を行う。ジェスチャーでのカーソル移動後、すぐに視線でのカーソル移動が発生するとクリックやドラッグ&ドロップの操作性が落ちるため、操作後0.5秒間は視線での移動は行わない。

3.2 視線取得部

今回、視線の取得には、高精度な視線計測が可能なTobii Eye Tracker 4Cを用いた。Tobii Eye Tracker 4Cでは、注視点がディスプレイ座標として出力される。この注視点をカーソル座標として更新することで視線でポインティングを行うことが可能である。しかしながら、出力される注視点は、前節で説明を行った固視微動の影響を受けている。そのため、出力値をそのままカーソル座標とするとカーソルが振動してしまう。そこで、ローパスフィルタの一種である指数移動平均 (Exponential Moving Average; EMA) フィルタを用いることで

この振動を軽減する。

指数移動平均フィルタの式を(1)式に示す。指数移動平均フィルタはパラメータとして、平滑化係数 α をもつ。入力 $x[n]$ に対して出力 $y[n]$ は(1)式のように表される。

$$y[n] = \alpha \cdot y[n-1] + (1-\alpha) \cdot x[n] \quad (1)$$

平滑化係数 α は0～1の値を取り、値が大きいほど平滑化の度合いが強くなる。測定された視線経路に対して指数移動平均フィルタを適用した結果を図4に示す。図4(a)では固視微動の影響が大きいものに対して、図4(b)では、平滑化により固視微動の影響がほとんどなくなっていることがわかる。しかし、平滑化が強すぎると、視線移動が高速である利点が活かされない。そこで、2種類のフィルタを用意し、注視点とカーソルの距離によって適用するフィルタを変更することで、カーソルの振動を抑え、高速なポインティングを可能とする。

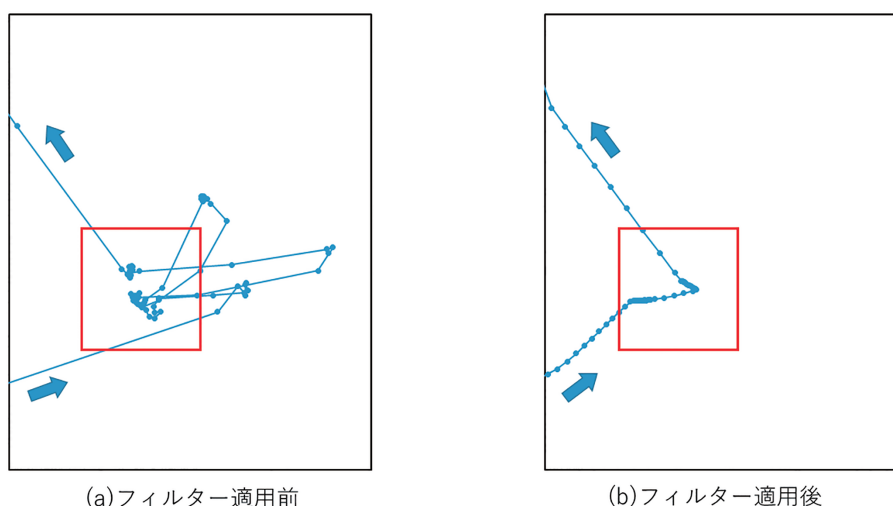


図4 視線経路の比較

3.3 フィードバック

本提案では、音と視覚の2種類のフィードバックがある。まず、音フィードバックでは、クリックなどの選択操作が発生するたび、対応した音を流す。次に、視覚フィードバックでは、視線でのポインティングとジェスチャーでのポインティングの状態をわかりやすくするため、円形のマスクを用意する。マスクの概要を図5に示す。これはカーソルと同時に移動し、その色によって現在の入力状態をユーザにフィードバックする。入力状態は三つあり、図5に示すように、視線でポインティングを行う状態は緑色、手で8方向のいずれかに移動を行う状態は赤色、手がLMCに認識されていない状態では移動を行わない黒色に変化する。

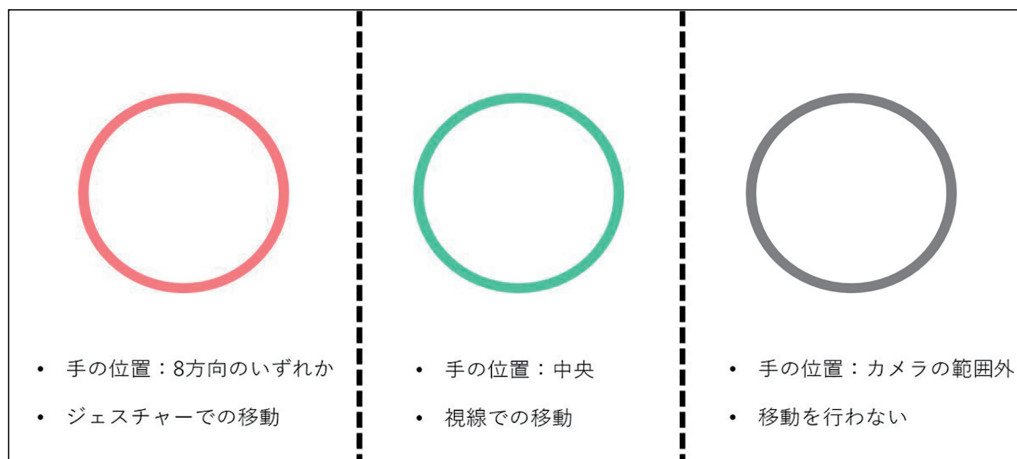


図5 マスク色

4 実験

本節ではタッチパッドとの比較実験について説明する。実験の被験者は8名、試行回数は1回、使用したディスプレイの大きさは14インチである。

4.1 実験方法

提案システムの評価を行うために、日常的によく選択されると思われる対象の選択をタッチパッドと提案システムで行った。被験者が行う試行をタスクと呼ぶ。タスク数は、タスク1～4が計測実験、タスク5がアンケートのみの実験で計5タスクである。計測実験では、タスク完了までの時間、ボタン以外を選択した回数を計測した。

タスク1～4の概要図を図6の(a)～(c)に、実験風景を図6の(d)に示す。タスク1・2では、ボタンのクリック操作を行った。(a)に示す1～9の番号順にボタンをクリックしていく。タスク1と2ではボタンの大きさが異なり、タスク1では80×80 pxのボタン、タスク2では40×40 pxのボタンのクリックを行った。タスク3では、ボタンの短距離でのドラッグ&ドロップを行った。(b)に示す中央の領域にはボタンが配置されており、このボタンを対象の領域へドラッグ&ドロップする。領域へのドロップに成功するとボタンは中央に戻る。この繰り返しを1～8の領域に対して順に行う。ボタンのサイズは40×40 px、領域のサイズは80×80 pxである。タスク4では、ボタンの長距離でのドラッグ&ドロップを行った。(c)に示す1～5の番号順にボタンを領域へドラッグ&ドロップする。ボタンと領域のサイズはタスク3と同じであるが、タスク3と違いボタンはドロップをされても初期地点に戻らずドロップ時の場所に留まる。タスク5では、Webブラウジングを以下の順に行った。

1. Google Chrome を選択する
2. ブックマークバーから Yahoo! ニュースを選択する

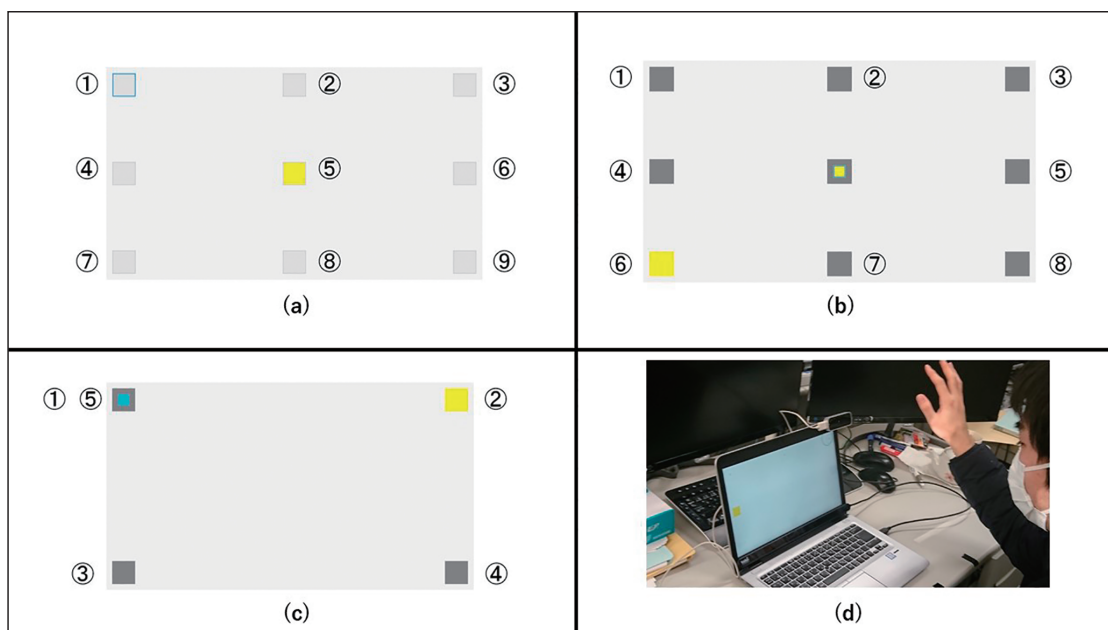


図6 タスク概要

3. 任意の記事を2件以上選択、閲覧、スクロールする
4. 閉じるボタンを選択しブラウザを閉じる

4.2 実験結果

4.1項で説明した評価実験およびアンケートを実施した。タスク1～4で測定した平均試行時間と平均誤操作回数を箱ひげ図で表したものを図7、図8に示す。なお、図7、図8において、ひげの下端は最小値、上端は最大値、箱の下端は25%、上端は75%の値を表している。また、箱内部の横線は中央値、×印は平均値、箱の外にある点は外れ値を表す。システムに関するアンケートを表1、2に示す。選択式アンケートの回答結果はどの選択肢に対して、何人が回答したかを示している。5択の質問は5段階評価として扱い、回答は1～5点に割

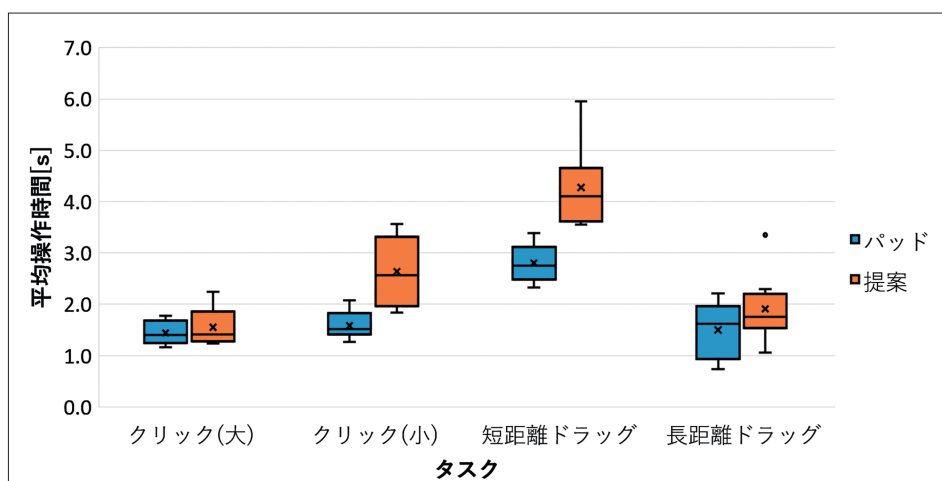


図7 平均試行時間

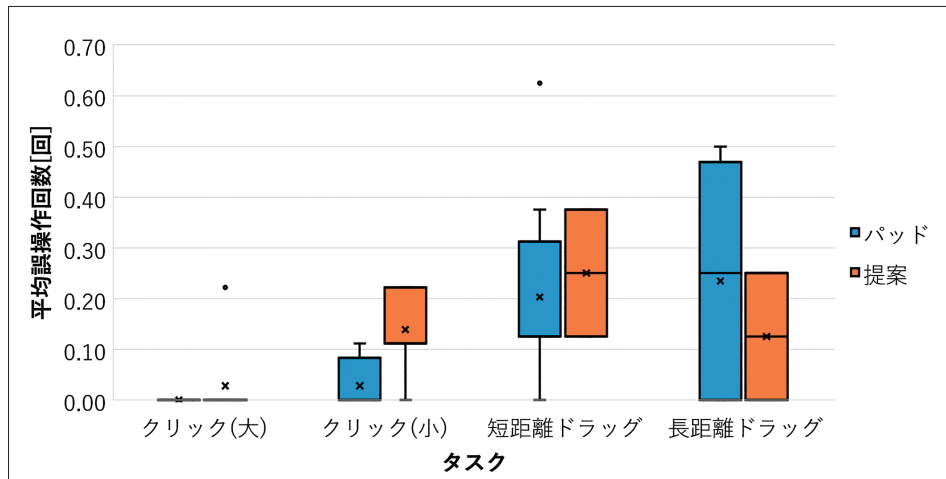


図8 平均誤操作回数

り当てを行い、最頻値を赤色の文字で示す。また、有効数字2桁の平均値を示す。

表1 提案システムの機能に関する回答結果

質問番号	質問	1点	2点	3点	4点	5点	平均値
A1.	クリックはどうでしたか?	0	0	0	3	5	4.6
A2.	ドラッグはどうでしたか?	0	0	2	3	3	4.1
A3.	ドロップ操作はどうでしたか?	0	0	2	2	4	4.3
A4.	視線でのカーソル移動はどうでしたか?	0	0	1	4	3	4.3
A5.	手のジェスチャーでのカーソル移動はどうでしたか?	0	0	0	5	3	4.4
A6.	提案システムの使用で目は疲れましたか?	2	2	1	3	0	2.6
A7.	提案システムの使用で手は疲れましたか?	1	1	1	4	1	3.4
A8.	総合的に見て、提案システムの操作性はどうでしたか?	0	0	1	5	2	4.1
A9.	提案システムを日常的に使用したいと思いましたか?	0	4	1	3	0	2.9

表2 提案システムとタッチパッドの比較

質問番号	質問	はい	いいえ
A10.	タッチパッドよりクリックしやすかったですか?	3	5
A11.	タッチパッドよりドラッグ&ドロップしやすかったですか?	6	2

4.3 考察

まず、タッチパッドとの比較を行う。図7に示すグラフから、大サイズのボタンクリックでは、タッチパッドと同程度の操作時間、誤操作回数で操作可能であることがわかった。これは、フィルターの改良により視線の俊敏性を保ちながら注視点の振動を軽減したことで、ジェスチャーによる微調整なしにクリックが可能であるためだと考えられる。一方、小サイズのボタンクリックでは、操作時間は長く、誤操作回数は多かった。この操作時間が長い原因は、ジェスチャー操作後に発生する0.5秒のカーソル停留時間によるものだと考えられる。小サイズのボタンでは注視点と選択対象の間にずれが生じやすいため、ジェスチャーの微調整を行う必要があり、カーソル停留によって選択時間が長くなるものと考えられる。長距離

のドラッグにおいて、操作時間の中央値、平均値はあまり変わらず、誤操作回数はタッチパッドよりも低かった。誤操作回数が低い理由として、長距離の移動の際、タッチパッドでは面積が限られているため、カーソルがオーバーシュートしやすいのに対し、提案システムでは視線により1動作で選択対象付近へのカーソルが移動が可能であるためだと考えられる。

次にアンケートの結果について考察を行う。クリックに関する質問では、低評価への回答はなく、平均値は4点を超えた。このため、パソコン操作で日常的に選択するアイコンであれば、クリック操作が可能であると考えられる。ドラッグに関する質問では、低評価の回答はなく、平均値は点を超えたが、記述回答などからスクロールがしづらいなどの意見があった。

次に改良システムの操作性の質問では、低評価への回答はなく、平均値もすべて4点を超えた。このため、実装機能に関しては問題なく使用できると考えられる。しかし、日常的な使用感の質問では、低評価への回答が多く、平均値も3点を下回った。原因としては、スクロールなどの実装機能が不足していたこと、普段は行わない目の操作やジェスチャー操作への疲労感が大きいといったことが考えられる。このうち疲労感はシステムを使っていくうちに慣れていき軽減されるものと考えられる。

5 まとめ

近年、センシング機器の多機能化、高性能化、低価格化により、ジェスチャーや音声などを利用した様々な入力機器が注目されている。このうち、非接触インターフェースである視線やジェスチャー入力でパソコン操作を行うシステムなども開発されており、今後、非接触型インターフェースは広く普及されると考えられる。

非接触の入力である視線入力には、サッカードと呼ばれる特性がある。このサッカードを利用して視線入力をポインティング操作に活かすことで、マウスやタッチパッドよりも高速なポインティングが可能となる。しかし、視線入力には、まばたきや固視微動といった不随意性のノイズや小さな選択対象の選択は困難であるという問題がある。

視線入力とは別の非接触型入力にジェスチャー入力が存在する。これは、身振り手振りでパソコンを操作するという手法である。ジェスチャー入力は、クリックなどの選択操作は得意であるが、操作回数が増えるにつれて手への負担が大きくなる。そこで、カーソル付近までの大まかな移動を視線で行い、微調整やクリックなどの選択操作をジェスチャーで行う。これにより、視線の俊敏性を活かしつつ、小さな対象への選択を行えるシステムを実装し、評価のためタッチパッドとの比較実験を行った。

その結果、80 px × 80 px のボタンと長距離のドラッグのタスクでは、タッチパッドと同程度の操作時間、誤操作回数であった。しかし、40 × 40 px のボタンクリック、短距離ドラッグのタスクではタッチパッドに比べ、操作時間、誤操作回数ともに大きかった。アンケート

の結果では、クリック操作は平均4点を超えた。しかし、ドラッグ操作時のスクロールバーの操作では平均4点を下回った。また、日常的な使用感では、低評価への回答が多く、平均3点を下回る結果となった。これは、スクロールがやりづらい、戻るといったショートカット操作がないことが原因だと考えられる。

参考文献

- 1) 金子寛彦, “知っておきたいキーワード 固視微動”, 映像情報メディア学会誌 Vol.63, No.11 (2009): 1538-1539
- 2) 橋村勝, 飯塚博実, 李軍, “特集3 人間工学のための計測手法 第4部: 生体電気現象その他の計測と解析(2) —眼球運動の計測—”, 人間工学 Vol.51, No.6 (2015): 406-410
- 3) 大和正武, et al, “一般的な GUI に適した視線・マウス併用型ターゲット選択方式”, 情報処理学会論文誌42.6 (2001): 1320-1329
- 4) Leap Motion : <https://www.leapmotion.com/ja/>