

HCI研究における評価実験用ビデオゲームの要件探究とオープンビデオゲームライブラリを用いたケーススタディ

岡 拓也^{1,a)} 川島 拓也¹ 築瀬 洋平² 渡邊 恵太³

概要: 我々は研究利用しやすいオープンビデオゲームライブラリを研究開発している。本論文はゲーム体験を記録するアーカイブ機能と、ゲームのビジュアルとパラメータの編集機能を追加し、評価実験のサポートについて紹介する。さらに、先行研究を事例に本ビデオゲームを用いた場合のケーススタディを行い、評価実験用ゲームとしての有用性を検証した。

キーワード: ビデオゲーム, 評価, コントローラ, ゲーム体験, オープンビデオゲームライブラリ

1. はじめに

ビデオゲームは映画や音楽と同じ用に日常的に楽しむコンテンツの1つとして広く浸透している。そしてビデオゲームは研究対象にもなっており、より優れたゲーム設計方法や利活用の研究も進んでいる。HCI分野においても、ビデオゲームに関連する研究は多数行われている [1]。

一方でビデオゲームに関連する研究を行う場合、コンテンツの取り扱いにいくつか問題がある。実験上の親しみやすさや理解のしやすさから、知名度の高い市販のビデオゲームを研究に取り込みたいが、ソースコードは公開されていないため編集や改変はできない。またそのキャラクターを使う際には著作権の配慮もしなければならない。ビデオゲームメーカーは研究目的での使用は想定しておらず、個別規定は困難であるため一般的な利用範囲や著作権についてしか明記していない。そのため研究者は市販ゲームを利用して研究することは、利用規定の不確実さや疑問からなる不安を払拭できない。

そこで、研究者はビデオゲームを自作することが多くある。例えば渡辺ら [2] は、顔の動きがアクションゲームに与える影響を明らかにするため、自作のアクションゲーム (図 1) を用いて実験を行った。その他にも Han ら [3] や萩原ら [4] の研究においても自作したゲームが使われている。研究用に自作したゲームの多くは、市販ゲームと比較するとゲーム性 (駆け引きの設計など) や、グラフィック表現や

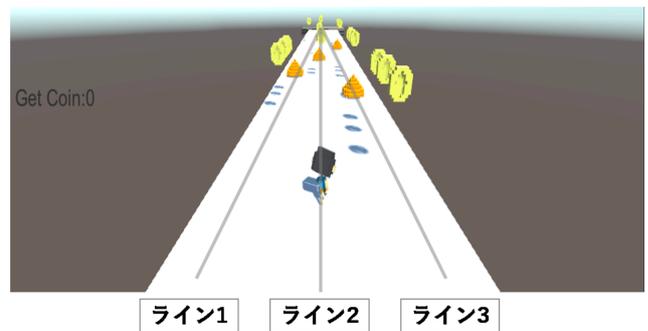


図 1 渡辺ら [2] が開発したアクションゲーム

音楽の質が簡素になる。これは制作コストや開発スキルの問題からなる。そしてそれを用いて実験すると、市販ゲームに慣れている実験参加者は、品質の違いや悪さから適切な印象や感想を得ることが難しくなることがある。これはデモンストレーションや論文発表においても同様である。

また各研究者が自作したビデオゲームを評価実験に用いると、研究間の比較がしにくい。例えば変形可能なゲームコントローラのユーザビリティを評価する福永ら [5] の研究と Shigeyama ら [6] の研究がある。2つの研究で提案されたゲームコントローラは、ゲーム内で剣を振る際に適した形状に変形可能である。このような研究目的やシステムの提案動機が似ている研究間においても、利用ゲームの違いから実験結果の比較がしにくい。すなわち、ビデオゲーム用いた研究はコンテンツ次第で研究の評価が左右されてしまう可能性がある。

そこで我々は、ビデオゲームを用いた研究において誰もが利用可能なオープンビデオゲームライブラリ (図 2) を提案した [7]。本論文では、これまでに集めた研究者からの要

¹ 明治大学大学院 先端数理科学研究科
先端メディアサイエンス専攻
² ユニティ・テクノロジーズ・ジャパン株式会社
³ 明治大学 総合数理学部 先端メディアサイエンス学科
a) cs212022@meiji.ac.jp

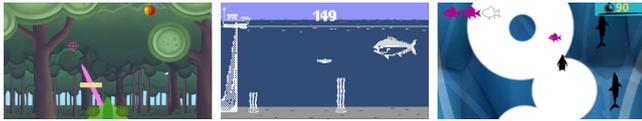


図 2 著者らがこれまでに開発したオープンビデオゲームライブラリ [7]

望や、ゲームを用いた HCI 研究の調査結果を基にして評価実験用ビデオゲームの要件を整理し、評価実験で利用しやすい機能を追加する。また、前論文 [7] ではオープンビデオゲームライブラリの利用方法や、対象研究について言及していなかったため、ケーススタディを行い利用しやすい環境について考察する。

2. 関連研究

この章では HCI 研究における評価実験用ゲームの要件を整理するために、2.1 で既存研究のゲームを用いた実験の評価対象と評価手法を調査する。そして、2.2 で既存研究が使用している評価実験用ゲームを調査し、需要と問題点を整理する。

2.1 ゲームを用いた実験の評価対象と評価手法

2.1.1 ゲーム体験の評価

(1) ゲーム体験の定量的評価

新たなゲームコントローラや体験手法を開発し、ゲーム体験を向上させる試みが多くなされている。例えば大山ら [8] は、快適な作業空間を提供する自律変形壁型ディスプレイを用いた体験手法を開発し、ゲーム体験の没入感や臨場感を高めた。一方でゲーム体験の客観的な評価は難しい。そのため The Player Experience of Need Satisfaction (PENS)[9] や The Game Experience Questionnaire(GEQ)[10] といったアンケートを用いて評価することがある。PENS は、不確かとされてきたゲーム体験の楽しさとプレイヤーの満足度をより詳細で正確に評価するために開発された。GEQ は、没入感、フロー体験、有能感、ポジティブ感情、ネガティブ感情、緊張感、チャレンジ感の 7 要素からゲーム体験を評価する。アンケートの他に、視線情報を計測してプレイ中の集中度を評価する研究もある [11]。

これらのゲーム体験の評価手法から、利用ゲームがゲーム体験の評価に影響を与える可能性が大いにあると考えられる。例えば、GEQ で評価する没入感や、ゲーム内の駆け引きや、グラフィック表現と音が相まった演出が影響するだろう。視線情報を計測する場合も同様に、グラフィック表現などが影響することが考えられる。つまり、実験用ゲームであっても、ゲーム体験を評価する場合は一般的に遊ばれている市販ゲームと同等の完成度のゲームを使う必要がある。そのため、本ゲームライブラリも完成度の高さが求められる。

(2) ゲーム体験の定性的評価

HCI 研究では人とコンピュータのインタラクションを扱うため、人を対象にした評価を行うことが求められる。ゲームを用いた研究も同様であるが、ゲーム体験の定まった定性的評価手法はない。Diah ら [12] は、子供を対象とした教育ゲームのユーザビリティを、観察法を用いて評価した。実験者はゲームをしている子供たちを観察しながら、全ての子供たちの言動や表情を記録した。ゲーム画面、ユーザが評価対象を操作している様子、その際に発したコメントを記録するためにビデオカメラを利用して記録した。

2.1.1 で定量的なゲーム体験の評価を紹介したが、ゲーム体験が完全に定量化されているわけではない。そのため定性的評価は必要だが、客観的な評価をする場合観察対象が多く、Diah らのように記録システムを構築する必要がある。本ゲームライブラリは貢献の 1 つとして研究者の負担を減らすことを掲げているため、このような記録システムを構築し定性的評価がしやすい環境作りをする。

2.1.2 ゲームコントローラの評価

ゲームと人の接点となるゲームコントローラに関する研究は多くなされている。HCI 分野においては、開発したゲームコントローラを用いた評価実験をすることも同様に多いが、評価手法は定まっておらず、多様な評価手法が用いられている。Brown ら [13] は、McNamara ら [14] の提唱するインタラクションの理論的フレームワークを用いて、ゲームコントローラの機能性、ユーザビリティ、経験の 3 側面を並行して評価する必要があることを示している。機能性はインタラクションの技術的側面のことで、ロギングソフトウェアを用いてコントローラが生み出す出力の量と範囲を、認識可能な入力と比較することで評価する。ユーザビリティはインタラクション自体のダイナミズムを見るもので、効率的、効果的で満足度のいくものかどうかの評価項目である。効率はコントローラを使用するために必要な精神的な努力を自己申告による Subjective Mental Effort Questionnaire (SMEQ) [15] を用いて測った。効果はタスクを正確に速くこなせているのか、ゲームスコアを用いて測った。Brown らのケーススタディではレースゲームを評価に利用したため、ラップタイムが使われた。ユーザがコントローラを使用してより速くラップを完了することができれば、より効果的なインタラクションとなる。満足度は、電子消費財に対するユーザの満足度を評価するための標準的な指標である Consumer Product Questionnaire (CPQ) [16] を用いて測定された。経験は、インタラクションの純粋に人間的な側面のことで、Critical Incidents Technique (CIT)[17] を用いて、ユーザの経験を記述する質的データが収集された。

Brown らのケーススタディでは、タスクを正確に速くこなせているのかをレースゲームのラップタイムで評価した。このパラメータを評価に利用した理由の記述は無かったが、タスクの明確さと推測する。評価対象や実験参加者

によってばらけてしまう不明瞭なタスクでは、適切な評価ができないためである。本ゲームライブラリも明確で分かりやすいタスクを設定する必要がある。

2.2 既存評価実験用ゲームからわかる需要と問題点

ゲームを研究利用する場合、認知度や完成度の高い市販のゲームの需要が高い。しかし、著作権の問題から編集することができず、評価実験用ゲームを実験条件に適応できない可能性がある。そこで、オープンソースとして非公式に公開しているクローンゲームを利用することがある。Karakovskiy ら [18] は、任天堂の古典的なプラットフォームゲーム「スーパーマリオブラザーズ」*1のクローンゲームである、Infinite Mario Bros.*2をベースに The Mario AI benchmark を開発した。Bellemare ら [19] は、Atari2600*3のエミュレータゲームである Stella*4を基に、Arcade Learning Environment を開発し、何百もの Atari2600 のゲーム環境を提供している。これらの研究用ゲームを開発した研究は引用数が多く、市販ゲームの認知度や完成度の高さと、実験条件に適応するための編集機能を併せ持つゲームの需要が高いことが分かる。

一方で、この両方を併せ持つゲームは少ないため、利用ゲームを妥協して選定することもあるが、問題が生じることがある。牛尾ら [20] は、情動制御の手法として温度提示を可能としたコントローラを作成し、既存のゲームの動きを認識して状況に応じて温度提示を行うシステムを開発した。自作ゲームの場合、ゲームの動きや状況に応じてインタラクティブに温度提示を行うことは容易だが、市販ゲームの場合難しい。この研究も市販ゲームを使用しているため、以下のような煩雑で間接的な方法で状況に応じた温度提示を行っている。起動後にコンテンツの制限時間が表示されている部分を画面の座標で指定する。その後、Python で OCR エンジンを利用可能にするモジュールの PyOCR と OCR エンジンの Tesseract*5を用いて、その部分の読み取りができなくなるまで値を読み取り続ける。このような方法をできる限り避け、ゲーム内で記録される特定の値を容易で直接的に取得するために、ゲームの編集機能が必要と考える。

ゲームインタフェースに関する研究の場合、インタフェースの編集機能が必要である。Iacovides ら [21] は、ゲームインタフェースをプレイヤキャラクターが見れるダイジェスティックなインタフェースと、プレイヤのみが見れるノンダイジェスティックなインタフェースに分けて考えた。ノンダイジェスティックなインタフェースは、プレイヤに有益な情報を提供する一方で、ゲームのストーリー

性を阻害する可能性があると考えた。人気という理由から Battlefield3*6が実験に用いられ、ゲームのインタフェースはコンソールに特定のコマンドを入力して変更された。変更パターンは、ノンダイジェスティックなインタフェースの一括非表示に限られ、著者が本来検証したい変更パターンが市販ゲームの編集の自由度の低さから検証できなかった可能性がある。

これらを踏まえ、以下の理由から本ゲームライブラリには編集機能が必要と考える。

- ゲーム内で記録される特定の値を容易で直接的に取得するため。
- 評価実験用ゲームを実験条件に柔軟に適応させるため。

3. 容易で適切な評価実験をするための機能開発

本ゲームライブラリを提案してから、様々な研究者の要望の収集や、ゲームを評価実験に用いた研究の調査をした。そこで得られた知見から、評価実験用ビデオゲームの要件を 3.1 に記す。さらに、要件の中で特に優先度の高い実験の録画機能とゲーム編集機能を開発した。ゲームパラメータの編集機能は前論文 [7] で既に開発したため、本研究ではゲームインタフェースの編集機能を開発した。

3.1 評価実験用ビデオゲームの要件

(1) パラメータとインタフェースの編集

2.2 で紹介した Iacovides ら [21] の研究のように、市販ゲームを用いる場合は編集できる要素が少なく評価実験に制約が生じる。そこで、編集機能があることで多様な実験条件に適応できる。また、本ゲームライブラリは多くの研究者に使ってもらい標準的な研究用ビデオゲームを目標にしているため、汎用性を高める機能は必要である。

(2) パラメータの出力

2.1.2 と 2.2 で紹介したように、実験のゲームプレイで記録されたゲーム内パラメータが評価分析に使われることがある。市販ゲームの場合、コードにアクセスすることができないため、間接的で煩雑な方法を取る必要がある。しかし、パラメータの出力機能があることで容易に実験結果を評価分析に用いることができる。

(3) 実験の録画

2.1.1 で紹介した Diah ら [12] の研究のように、画面上で起きているゲーム内容と並行して実験時の参加者の言動をビデオカメラで記録することがある。ゲーム体験の評価は、観察対象が多くリアルタイムで済ませることが難しい。そのため、録画機能を用いてプレイ後に分析することで、負担が少なく、正確な定性評価ができると考える。

*1 <https://www.nintendo.co.jp/software/smb1/index.html>

*2 <https://www.smbgames.be/infinite-mario.php>

*3 <https://ja.wikipedia.org/wiki/Atari2600>

*4 <https://stella-emu.github.io/>

*5 <https://github.com/tesseract-ocr/tesseract>

*6 <https://www.ea.com/ja-jp/games/battlefield/battlefield-3>

(4) ルールが分かりやすく、市販ゲームと同等の完成度のゲーム

2.1.1 で記したように、利用ゲームがゲーム体験の評価に影響を及ぼす可能性があるため、市販ゲームと同等の完成度のゲームを実験に利用する必要がある。さらに、実験での参加者の学習コストや理解度を考慮し、分かりやすいルールが好ましい。

(5) ゲーム内乱数の再現

エンタテインメントコンピューティング 2021^{*7}にて研究者から寄せられた要望に、ゲーム内乱数の再現機能がある。慣れ防止や開発の都合から、ゲームでは乱数を使うことが多いが、実験において参加者間で実験条件を厳密に揃える場合、ゲーム内乱数の再現機能が必要である。

3.2 実験を記録するアーカイブ機能

この機能は、実験でゲームをプレイする参加者の様子を記録し、評価分析に活用することを目的に開発した。ゲーム画面と参加者の様子を映した映像がまとめて出力され、一度に観察できる(図3)。定性的な評価分析をする際に、出力された映像を利用することで、字面だけでは分からない、回答の意図を読み取れる可能性がある。例えば、参加者の表情から、提案手法の良い点・悪い点が見つかりやすくなったり、手元の入力の様子を映した映像と、入力を可視化したユーザインタフェースを見比べることで、入力データには現れなかったエラーを見つけ出すことができる。ゲームを用いた評価実験は、参加者のあらゆる言動とゲーム画面が観察対象だが、リアルタイムで全てを記録するのは難しい。そのため、実験の内容を記録する機能に対して一定の需要があると考えられる。しかし、ゲームを用いた実験の記録システムは標準化されていない。そのため、研究者がこのような実験記録機能を自作する必要がある。Tanら[22]は、ビデオゲームのプレイ中に収集された声のデータと生理学的データを組み合わせてゲーム体験を評価する手法を提案した。この評価手法で利用された実験記録機能(図4)を用いると、ゲーム画面、参加者の上半身全体の様子、参加者の表情、視覚化された生体情報の4つを一度

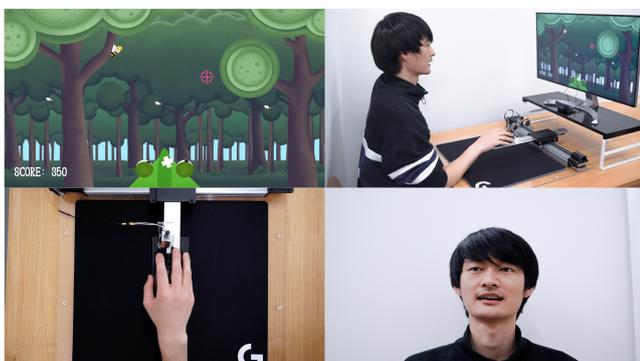


図3 アーカイブ機能を用いた実験のモニタリング画面



図4 Tanら[22]が評価実験に利用した実験モニタ画面



図5 ゲームインタフェースの調整画面

に観察できる。このシステムは公開されていないため、他の研究者が利用することはない。本研究のアーカイブ機能は、オープンビデオゲームライブラリの機能の1つとしてオープンに公開する。

3.3 ゲームインタフェースの編集機能

ゲームインタフェースに焦点を当て、ゲーム体験への影響を調査する研究がある[21]。しかし、市販ゲームを利用する場合、研究利用は想定されていないためユーザが編集できる範囲は限られている。つまり、研究者が本来検証したいことが市販ゲームの編集の自由度の低さから阻害される可能性がある。このような問題を解決するため、ゲームのインタフェースを自由に変更することや、表示の有無を選択することができるゲームインタフェースの編集機能(図5)を開発する。また、著者らはこれまでにパラメータの編集機能を開発したため、これと統合させる。

4. オープンビデオゲームの利用とケーススタディ

研究者の評価実験用ゲームの要件を探究し、評価実験に使いやすい機能を備えた研究用ゲームを開発した。しかし、実際に研究者が求める評価実験用ゲームとして有用であるかは未検証である。そこで、既存のビデオゲームを用いた研究に対して、本ゲームライブラリの利用を想定したケーススタディを行うことで有用性を検証する。

^{*7} <http://ec2021.entcomp.org/>

4.1 ゲーム観戦におけるプレイヤー操作の追体験システムの定性的評価

佐藤ら [23] は、ゲーム観戦においてプレイヤーのゲーム体験をマウスの操作情報も含めて観戦者が追体験するシステムを提案した。プレイヤーのマウスの挙動を、観戦者用マウスでリアルタイムで再現することで自分もプレイしているように感じさせることが目的である。ユーザスタディではゲーム体験終了後、ゲーム観戦に関するインタビューを行い、自分も操作しているように感じるかを定性的に評価した。

このケーススタディでは、本ゲームライブラリのゲームと実験を記録するアーカイブ機能を用いて追体験システムの定性的評価を行う。佐藤らの実験でシューティングゲームが利用されていたことから、Hunter-Chameleon(図7)を評価実験用ゲームとして用いる。このビデオゲームは、カメレオンの舌でターゲットを撃ち抜く Shoot ' Em Up ゲームである。プレイヤーの目標は、制限時間内に次々と出現するハエや果物といったエサをできるだけ多く舌で入手し、スコアを伸ばすことである。この実験では、自分も操作しているように感じるかを評価するため、思ったことや感じたことを発話するシンクアラウド法を行う。さらに、アーカイブ機能を利用してレトロスペクティブレポート法 [24] を併せて行い、参加者自身に実験時の発話内容を振り返ってもらう。

4.2 評価実験から分かった本ゲームライブラリの有用性

実際に実験を記録し(図3)、参加者にはプレイ後に映像を見ながら自身のプレイを振り返ってもらった。その際の発話内容から、本ゲームライブラリとアーカイブ機能が有用であったと言える点を記す。

参加者が実験を振り返る際、「実況(発話)に現れているか分からないけど」と前置きをした上で実験時の心境を語る場面があった。その後、アーカイブ機能を用いて参加者と発話内容を振り返ることで、参加者の言っているシーンのタイミングや意図を理解することができた。また、参加者が実験時のことを思い出そうとしている際に、「自分のコメント振り返ればわかるかも」と発言した。このシーンから、参加者自身も実験時の記憶をアーカイブ機能を用いて思い出すことができることを確認した。

また、評価対象のシステムを開発した研究者に本ゲームライブラリを用いて評価実験をした感想を聞いたところ、「複数視点を同時に見れるアーカイブ機能が嬉しい」と評価してもらった。本ケーススタディのように、本来ゲーム体験の定性的評価は観察対象の多く、評価が難しい。しかし、本ゲームライブラリを使うことで、実験者に記録の負担をかけることなく定性評価ができることを示す好例となった。さらに、ゲームシステムに関して、「ルールが分かりやすく実験に使いやすい」と評価してもらった。これ



図6 著者が開発した Shoot ' Em Up ゲーム, Hunter-Chameleon は、動物のメタファーを用いて、ルール説明せずとも理解できるようなゲーム設計にしたことが効果的であったと考えられる。

4.3 本ゲームライブラリのベンチマーク性を利用した、複数入力デバイスの相対評価

Natapov ら [25] は、ファーストパーソン・シューティングゲームにおけるユーザのパフォーマンスを、プロトタイプのトラックボールコントローラー、一般的なゲームコントローラー、キーボードとマウスの計3種類の入力手法で比較した。ゲーム用入力デバイスの評価手法は様々あるが、断片的ではなく、一般利用を想定した総合的な評価が可能である点から、Brown ら [13] の評価手法を基に評価実験を行う。この手法は、入力デバイスを経験性、機能性、ユーザビリティの3つの側面から評価する。

経験性は、Critical Incidents Technique (CIT)[4]を用いて、ユーザーの経験を記述する質的データを収集して評価されるが、本ゲームライブラリの有用性検証につながらないため行わない。

機能性は、入力デバイスが生み出す出力の量と範囲を、認識可能な入力と比較することで行われる。これは入力デバイスの情報を収集して行われるが、本ゲームライブラリの有用性検証につながらないため行わない。

ユーザビリティは、効果、効率、満足度の3つの評価基準がある。効果的かどうかの評価基準は、タスクを正確に速くこなしているかであり、ゲームスコアを用いて測定する。効率的かどうかは、コントローラを使用するために必要な精神的な努力を自己申告による Subjective Mental Effort Questionnaire (SMEQ) [1]を用いて測定するが、経験性と同様に本ゲームライブラリの有用性検証につながらないため行わない。満足度は Consumer Product Questionnaire (CPQ) [16]を用いて測定するが、効率性と同様の理由で測定はしない。

また、入力デバイスの研究において、他の入力デバイスとの比較はあまり行われていない。この理由として、評価実験に利用できる標準的なゲームが無く、実験条件が研究



図 7 入力手法評価実験で評価する入力デバイス. Razer BASILISK^{*14}(左), Xbox Wireless Controller^{*15}(中央), LeapMotion^{*16}(右)

によって異なることが考えられる. そこで, 本ゲームを用いて同じ実験条件下で, 複数入力デバイスを評価し, 入力デバイスの相対的評価を行う. 本ゲームライブラリにはパラメータ編集機能があり, 研究者が自身の研究条件に合わせて変更できるが, 本ケーススタディではベンチマークとして利用するため, パラメータを統一する. インタフェース編集機能によって見た目の変更も可能であるが, 同様の理由から統一する. 本ケーススタディで評価する入力デバイスは, ゲーミングマウスとして評価の高い Razer の BASILISK^{*8}と, 市販のゲームコントローラとして代表的な Xbox Wireless Controller^{*9}, ハンドトラッキングによる入力可能な Leap Motion^{*10}の3つとする. 以後 Razer の BASILISK を BASILISK, Xbox Wireless Controller を Xbox と表記する.

4.3.1 評価実験

実験は3種類の入力手法で Hunter-Cameroon を高得点を目指してプレイしてもらう. 参加者は20代の男女6名で, 各入力手法で2回ずつプレイし, 平均スコアを利用する. 持ち越し効果を防ぐため, カウンターバランスを行う. 入力手法が3種類のため, 全6通りの実験順序を6人の参加者各々に適用する. 実験手順は, ゲームの説明をした後, 1種類目の入力手法を説明し, 好きなだけ練習時間を設けた. 参加者が練習が十分であると報告された後, 1分間の実験プレイを2回行う. 他の入力手法も同様の手順で実験を行う. 実験で使用したゲームのパラメータは, りんごの出現頻度が5秒毎, 蜂は7秒毎, ハエは1秒毎. スピードは全て0.05とした.

実験から表1のスコアを取得した. 一元配置分散分析法を用いて, これらの結果の統計的有意差を確認できる. これに伴い, どの入力手法間に有意差があるかを調べるために, Tukey 法による多重比較を行う. その結果, BASILISK と Leap Motion の間と Xbox と Leap Motion 間で有意差を確認した(表2. このことから, ユーザビリティの効果的であ

*8 <https://www2.razer.com/jp-jp/gaming-mice/razer-basilisk>

*9 <https://accessories.trafficmanager.net/accessories/ja-jp/products/gaming/xbox-one-controller-cable-for-windows/7mn-00005>

*10 <https://www.ultraleap.com/product/leap-motion-controller/>

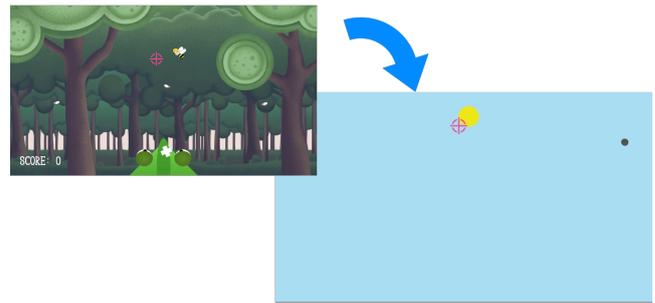


図 8 インタフェース編集機能を用いた Hunter-Chameleon のインタフェース変更例

るかという評価基準において, Leap Motion は BASILISK と Xbox Controller に劣っていることが分かる.

表 1 3種類の入力手法で記録されたスコア

	A	B	C	D	E	F
BASILISK	6700	3625	6800	6550	6200	6375
Xbox	4575	6125	4275	2900	2900	5600
Leap Motion	3100	175	1325	1500	0	325

表 2 スコアの多重分析をした結果

	BASILISK	Xbox	Leap Motion
BASILISK	-	3.25	9.82
Xbox	-	-	6.57
Leap Motion	-	-	-

4.3.2 評価実験から分かった本ゲームライブラリの有用性

このケーススタディは著者が本ゲームライブラリを利用する研究者になりきって行った. そのため, 本ゲームライブラリの客観的な評価とはならないが, 研究者視点で有用性についてまとめる.

評価に使ったパラメータは, 実験終了時の画面にあるダウンロードボタンを押して CSV ファイルとして出力した. 実験者にかかる負担は殆どなく, 評価実験とその後の分析を容易にする機能と言える.

今回の実験では, プレイタイムを60秒, 餌の出現頻度をハエが1秒毎, 蜂が7秒毎, りんごが5秒毎に設定している. 実験条件をパラメータ編集画面の GUI 操作のみで設定することができるので, 研究に適応させやすいゲームと言える.

4.4 ゲームインタフェースのゲーム体験への影響評価

Iacovides ら [21] は, ゲームインタフェースがもたらすプレイヤーの感じ方の違いがゲーム体験にどのように影響するのかを評価した. このケーススタディでもゲームインタフェースを編集機能を使い変更する. 低コストであらゆるケースを評価でき(図8), 特定のユーザインタフェースを消すことや, 別の画像に差し替えることができる. このケーススタディではゲームインタフェースを変更できることの確認までを行い, インタフェースの違いによるゲーム体験への影響の調査は行わなかった.

4.5 本ゲームライブラリの有用性

4.4 のケーススタディでは評価実験は行っていないが、様々なゲームインタフェースに変更してゲームをプレイできることを確認した。4.3 と同様に、本ゲームライブラリの客観的な評価とはならないが、研究者視点で有用性についてまとめる。

編集の自由度が高く、様々なパターンを試すことができた。画像を指定することで、背景、エサ、照準を合わせるレティクルの見た目を変更できた。また、サイズ調整もできるため、難易度調整もできる。GUI 操作で変更が可能なので、準備が容易であった。

5. 議論

本稿では、HCI 研究における研究用ゲームの要件を基にした機能の開発と、ケーススタディを行った。探究した研究用ゲームの要件は、妥当であり、十分なものであるのか議論する。

また、本ゲームライブラリは、ゲームを用いた研究において標準的に使われるデファクトスタンダードを目指す一方で、1つ1つの研究において利用される有用性を保つ必要がある。このバランスの観点からも議論をする。

5.1 要件は妥当かつ十分であるか

ゲームパラメータやインタフェースの編集機能は、様々な研究条件に適応させるために必要不可欠な機能である。しかし、より多くの実験条件に適応させるべく、編集の自由度を高めると扱いづらくなる。とはいえ、Hunter-Chameleon の自由度は低いため、調整可能パラメータを増やす必要があると考える。自由度を高めるためにはソースコードを公開することも検討する。

プレイ結果のパラメータ出力も評価分析を容易にする有用な機能である。しかし、ゲーム画面に表示され、容易に取得できるパラメータのみを分析に利用する場合、このような機能は必要ない。もちろん画面に表示される情報では取得できないパラメータも多くあるが、そのようなパラメータがどこまで必要とされているか調査できていない。分析に利用されるパラメータの調査を踏まえ、出力機能の有用性の検証を続ける必要がある。

実験の様子を記録するアーカイブ機能は、多くの研究の評価実験で有用である。ケーススタディでは3視点のカメラ映像を表示したが、入力や生体情報の可視化インタフェースや、ゲーム情報を表示すべき場合も考えられる。そのため、研究者が表示コンテンツを選択し、出力するとさらに多くの研究者が使いやすいものになると考える。

5.2 ベンチマークとしての公平性

本ゲームライブラリは研究領域において標準的に使われることを目指し、ゲームシステムの設計や機能開発をして

きた。標準的に使われることで、ベンチマークゲームとして、研究間の比較を促進させたいと考えている。4.3 では、複数入力デバイス間のベンチマークとして入力デバイスの評価をした。しかし、本ゲームライブラリは評価対象である入力デバイスと独立した関係ではなく、完全な公平性を保つことができていない。さらに、入力の仕組みが根本的に異なる入力デバイス間では、制御のマッピングが操作性に大きな影響を与え、実験結果を入力デバイスの優劣と結びつけるのは公平性に欠ける。公平性のあるベンチマークとしてのゲーム設計は今後の課題である。

5.3 利用規約の設定

市販のゲームやオープンソースゲームは、著作権の問題があり、どこまでをどう利用していいのかが不明瞭なため、安心して利用できない。そこで、オープンビデオゲームライブラリは、研究利用が可能であることを明言し、研究者が使いやすいゲームを目指す。以下に本ゲームの利用用途や利用時の留意点をまとめた利用規定を明記する。この規定の内容や表現が妥当であるかは今後継続的に考察する。

5.3.1 利用用途

本ゲームライブラリは研究目的のあらゆる利用を認める。

5.3.2 利用時の注意点

本ゲームは、各研究者の実験条件に適応するため、ゲーム内パラメータを変更できる。そのため、一見同じゲームを実験利用しているように見えて、実験条件が異なる場合がある。同一のゲームを複数研究間で利用することの利点として、実験結果の比較が容易であることがある。しかし、ゲーム内パラメータが異なると、実験結果をそのまま比較することができない。そのため、各研究者が本ゲームを実験利用した際に、設定パラメータを明記することで汎用的でありつつも、再現性のあるゲームとなる。より多くの研究者に利用されるために、使用パラメータの記載は本ゲームを利用するための必須条件とはしないが、使用パラメータを含めた実験条件の共有を推奨する。

本ゲームライブラリを利用して得られた研究成果(論文、報告書、出版物、派生ソフトウェア)を公開する場合には、引用し、本ゲームライブラリを利用した旨を明記する。明記の方法は特に指定しないが、例えば、「本研究は、オープンビデオゲームライブラリを利用した。」というように書いてもらう。

6. おわりに

HCI 研究における評価実験用ビデオゲームの要件探究を基に機能の追加し、ケーススタディを踏まえて有用性の検証をした。評価実験用ビデオゲームの要件を満たすゲームライブラリを作成するために今後も要件の整理と、それを満たす機能の追加、改善をする必要がある。また、本稿では有用性の検証の一部を著者自身が行ったため客観性が欠

けている。そのため、今後はゲームを用いた研究を行う研究者に、本ゲームライブラリの評価をしてもらう。

参考文献

- [1] Marcus Carter, John Downs, Bjorn Nansen, Mitchell Harrop, and Martin Gibbs. Paradigms of Games Research in HCI: A Review of 10 Years of Research at CHI. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '14, p. 27–36, New York, NY, USA, 2014. Association for Computing Machinery.
- [2] 渡辺泰伎, 白石陽. アクションゲームにおける顔の動きを用いたゲーム操作の拡張に関する考察. Technical Report 19, 公立はこだて未来大学大学院システム情報科学研究科, 公立はこだて未来大学システム情報科学部, dec 2019.
- [3] Changyo Han, Ryo Takahashi, Yuchi Yahagi, and Takeshi Naemura. Pneumodule: Using inflatable pin arrays for reconfigurable physical controls on pressure-sensitive touch surfaces. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–14, New York, NY, USA, 2020. Association for Computing Machinery.
- [4] 萩原正宏, 高橋伸, 田中二郎. ダメージを演出するための鋭い感触を呈示する手法の検討と開発. Technical Report 6, 筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻, 筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻, 筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻, may 2011.
- [5] 健竜福永, 竜太石河, 亮文井上. 変形により入出力が可能なゲームコントローラのユーザビリティ改善. エンタテインメントコンピューティングシンポジウム 2019 論文集, 第 2019 巻, pp. 46–51, sep 2019.
- [6] Jotaro Shigeyama, Takeru Hashimoto, Shigeo Yoshida, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. Demonstration of Transcalibur: A VR Controller That Presents Various Shapes of Handheld Objects. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, p. 1–4, New York, NY, USA, 2019. Association for Computing Machinery.
- [7] 岡拓也, 川島拓也, 林大智, 渡邊恵太. 研究利用しやすく標準性を目指したビデオゲームの設計と開発. エンタテインメントコンピューティングシンポジウム論文集, pp. 181–186, 2021.
- [8] 大山貴史, 高嶋和毅, 浅利勇佑, エフドシャーリン, ソールグリーンバーク, 北村喜文. 自律移動・変形する壁型ディスプレイの設計と実装. 情報処理学会論文誌, Vol. 58, No. 5, pp. 1049–1060, may 2017.
- [9] Richard M Ryan, C Scott Rigby, and Andrew Przybylski. The motivational pull of video games: A self-determination theory approach. *Motivation and emotion*, Vol. 30, No. 4, pp. 344–360, 2006.
- [10] Wijnand IJsselstein, Yvonne De Kort, Karolien Poels, Audrius Jurjelionis, and Francesco Bellotti. Characterising and measuring user experiences in digital games. In *International conference on advances in computer entertainment technology*, Vol. 2, p. 27, 2007.
- [11] 山浦祐明, 中村聡史. 視線に追従するぼかしエフェクトがビデオゲームの体験に及ぼす影響の調査. Technical report, 一般社団法人情報処理学会, 2019.
- [12] Norizan Mat Diah, Marina Ismail, Suzana Ahmad, and Mohd Khairulnizam Md Dahari. Usability testing for educational computer game using observation method. In *2010 International Conference on Information Retrieval Knowledge Management (CAMP)*, pp. 157–161, 2010.
- [13] Michael Brown, Aidan Kehoe, Jurek Kirakowski, and Ian Pitt. *Beyond the Gamepad HCI and Game Controller Design and Evaluation*, pp. 263–285. Springer International Publishing, Cham, 2015.
- [14] Niamh McNamara and Jurek Kirakowski. Functionality, usability, and user experience: Three areas of concern. *Interactions*, Vol. 13, pp. 26–28, 11 2006.
- [15] Albert G. Arnold. Mental effort and evaluation of user-interfaces: A questionnaire approach. In *Proceedings of HCI International (the 8th International Conference on Human-Computer Interaction) on Human-Computer Interaction: Ergonomics and User Interfaces-Volume I - Volume I*, p. 1003–1007, USA, 1999. L. Erlbaum Associates Inc.
- [16] Niamh McNamara and Jurek Kirakowski. Measuring user-satisfaction with electronic consumer products: The consumer products questionnaire. *Int. J. Hum.-Comput. Stud.*, Vol. 69, No. 6, p. 375–386, jun 2011.
- [17] J. C. Flanagan. The critical incident technique. p. 327–358, 1954.
- [18] Sergey Karakovskiy and Julian Togelius. The Mario AI Benchmark and Competitions. *IEEE Transactions on Computational Intelligence and AI in Games*, Vol. 4, No. 1, pp. 55–67, 2012.
- [19] M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling. The arcade learning environment: An evaluation platform for general agents. *Journal of Artificial Intelligence Research*, Vol. 47, p. 253–279, Jun 2013.
- [20] 牛尾大翔, 水口充. 既存のゲームシステムを温冷覚での情動変化で支援するシステムの開発. エンタテインメントコンピューティングシンポジウム論文集, 第 2021 巻, pp. 379–380, aug 2021.
- [21] Ioanna Iacovides, Anna Cox, Richard Kennedy, Paul Cairns, and Charlene Jennett. Removing the hud: The impact of non-diegetic game elements and expertise on player involvement. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '15, p. 13–22, New York, NY, USA, 2015. Association for Computing Machinery.
- [22] Chek Tien Tan, Tuck Wah Leong, and Songjia Shen. Combining think-aloud and physiological data to understand video game experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, p. 381–390, New York, NY, USA, 2014. Association for Computing Machinery.
- [23] 佐藤大輔, 渡邊恵太. ゲーム観戦におけるプレイヤ操作の追体験システムの試作. エンタテインメントコンピューティングシンポジウム論文集, 第 2021 巻, pp. 400–404, aug 2021.
- [24] 福田忠彦. 人間工学ガイドー感性を科学する方法ー IATeX. サイエンス社, 2004.
- [25] Daniel Natapov and I. Scott MacKenzie. Gameplay evaluation of the trackball controller. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology*, Futureplay '10, p. 167–174, New York, NY, USA, 2010. Association for Computing Machinery.