

アドベンチャーゲームにおけるユーザ体験向上のための感情推定技術を用いた入力手法

大山 智弘^{1,a)} 川口 一画^{2,b)}

概要: アドベンチャーゲームとは、テキスト、画像、音声等により物語が提示され、物語中の選択肢をプレイヤーが選択することによりシナリオが分岐するデジタルゲームである。既存のアドベンチャーゲームの入力方法は、マウス等を用いて提示された選択肢のテキストを選択する方法が一般的である。本研究では、アドベンチャーゲームの没入感および印象の向上を目的として、プレイヤーによる主人公の演技を用いた入力手法を提案する。プレイヤーは、ゲーム中の選択場面において、提示された主人公のセリフおよび感情の選択肢を元に主人公の演技（セリフの発話および表情の作成）を行う。システムは、音声認識により取得した発話内容、およびカメラを用いて発話時の表情から推定した感情を基にシナリオ分岐を行う。提案手法がユーザ体験に与える影響について検証するため、選択にマウスを用いるマウス条件、音声認識を用いる発話条件、音声認識と表情検出を組み合わせた演技を用いる演技条件を比較する実験を行った。その結果、演技条件はマウス条件と比べ没入感の向上が示された。一方で、ゲームに対する印象は条件間で差が見られず、さらに演技条件では入力時の緊張の増加が示された。

1. はじめに

現代のアドベンチャーゲームは、グラフィック、音声とともにテキストでシナリオが表示され、シナリオの途中に分岐点が存在し、そこで選択肢のテキスト選択を行いシナリオが分岐する形式のものが大半である。入力もゲームパッド、マウスおよびキーボードを用いたもので、プレイヤーは単に画面に表示されたテキストを読み、選択するのみである。しかし、本研究では、この入力手法よりもさらに没入感および印象を向上させることのできる入力手法が存在すると考える。

ここでアドベンチャーゲームと類似するゲームとしてテーブルトップロールプレイングゲーム (TRPG) を説明する。TRPGとは、アドベンチャーゲームと同様に、プレイヤーの行動によってシナリオが分岐するゲームである。アドベンチャーゲームと違う点も複数存在するが、その中の1つに複数人のプレイヤーが自分の作成したキャラクターをシナリオに登場させ、キャラクターを演じながらやりとりをし、ゲームを進めることが挙げられる。TRPGでは、この演技により次のプレイヤーの振る舞いが変化し、物語が複雑に展開する。プレイヤーは演技により複雑に展開する物語

に、アドベンチャーゲームと比較し自身の選択が物語の展開を動かす体験をより強く感じ、より高い没入感を得ることができると考える。ここで、本研究ではキャラクターを演じるという点をアドベンチャーゲームに対し応用することで、没入感の向上を図る。

本研究では、アドベンチャーゲームの選択において、従来手法のボタン、マウスによる入力ではなく、表情からの感情推定を用いた入力手法を提案する。これにより、プレイヤーは従来手法と比較してゲーム内の主人公に感情移入を起こし、ユーザ体験が向上すると考える。

2. 関連研究

本章では、まず表情の作成に関する心理学的知見を述べる。その後、表情からの感情推定を用いた先行研究およびゲームに関する先行研究を述べ、本研究の位置づけを示す。

2.1 表情フィードバック仮説

本研究では、アドベンチャーゲームをプレイ中のプレイヤーが持つ感情を増幅させることにより、没入感および印象の向上につながると考え、人間の情動の発生に着目した。

情動の起源としては中枢起源説および末梢起源説に大別される。中枢起源説 [1,2] とは、情動の起源は脳等の中枢機関であり、刺激により情動が生まれるという考えに基づく説である。一方、末梢起源説 [3] とは、情動の起源は表

¹ 筑波大学情報理工学学位プログラム

² 筑波大学システム情報系

^{a)} ohyama@iplab.cs.tsukuba.ac.jp

^{b)} kawaguchi@cs.tsukuba.ac.jp

情等の末梢部であり、刺激により末梢部の反応が起こり、情動が生まれるとする説である。1884年にJames [3]は末梢起源説を提唱した。その後、1962年にTomkins [4]が末梢起源説のうち、情動の発生には顔の表情が重要な役割を持つことを強調した。これがきっかけとなり顔の表情に限られた研究、すなわち表情フィードバック仮説の研究が行われるようになった。代表的な研究が、Strackらによる、被験者に口でペンを咥えさせ、意図を悟られることなく喜びと悲しみの表情を作らせる実験を行った研究 [5]で、有効性が実証された。

本研究では、アドベンチャーゲームにより引き起こされる情動変化を表情作成により強化し、没入感および印象の向上につなげる。

2.2 表情からの感情推定を用いた先行研究

表情フィードバック仮説の知見を取り入れ、日常生活で笑顔を促す新たな鏡および冷蔵庫を作成した研究 [6]について述べる。この研究では、鏡および冷蔵庫に提案デバイスを取り付ける。このデバイスにより、鏡ではカメラによって撮影されたユーザが笑顔になると、音とともにLEDディスプレイで笑顔アイコンを提示する。冷蔵庫ではLEDディスプレイおよび音でのフィードバックに加え、笑顔にならないと扉が開かない仕様にした。提案システムを導入した鏡および冷蔵庫を一般家庭に設置し経過観察を行う実験の結果、提案システムが笑顔促進に効果的に働き、家族間のコミュニケーションにも良い方向に影響を与えることが分かった。ここで、これらの表情からの感情推定技術をゲームに用いることで同様のユーザ体験をもたらすことができると考えた。

2.3 表情からの感情推定を用いたゲーム

表情からの感情推定を入力に用いたゲームに関する先行研究について述べる。2008年にLankesらによって発表された、“EmoFlowers” [7]である。このゲームでは、プレイヤーは画面を見ながら笑顔を作り、画面内の花を咲かせることにより、スコアを稼ぐ。このゲームは、表情による入力を前提とされた、ゲームパッド、マウス、キーボード等の既存の入力手法では代替できないゲームである。実験の結果、92.4%の参加者がゲームに対してポジティブなユーザ体験を得た、63.8%の参加者がゲームから感情の影響を受けたと回答した。

同様に、表情からの感情推定をEndless Runnerと呼ばれるゲームジャンルに応用した研究 [8]を挙げる。この研究では、プレイヤーが障害物を避ける動きを表情に対応させ、遠隔地の2人で遊ぶ仕様にしたゲームを開発した。このゲームでは、遠隔地の2人のプレイヤーが、ボイスチャットを通して協力をし、同時に同じ表情をしなければ障害物を避けられないようになっている。このゲームをキーボー

ド操作と比較し、評価した。実験の結果、ユーザ体験の向上および共にプレイした相手に対する親近感の向上が見られた。一方で、入力に対してキーボードの方がゲームプレイがよいという意見も多く見られた。

以上、表情からの感情推定技術を用いたゲームの先行研究を述べた。しかし、本研究で用いるアドベンチャーゲームのジャンルに使用されたことはない。また、表情を入力する場合でも、発話との組み合わせ、演技という文脈で用いられてはいない。

3. システム構成

本章では、提案手法である、音声認識、および表情からの感情推定を入力として行えるアドベンチャーゲームシステムの設計を述べる。

3.1 設計指針

本研究の提案手法は、2.1節において述べた情動の起源に関する心理学的知見に基づき、アドベンチャーゲームにおいて没入感および印象を向上させるための手法である。また、実験設計上、従来手法と比較できる入力手法である必要がある。以上のことを踏まえ、提案手法の設計指針は以下の通りである。

- ゲーム内の選択肢に、主人公のセリフ、およびその際の主人公の感情を表示する。
- 画面に選択肢が表示されている時に、選択肢を読み上げて主人公のセリフを選択する。また、表情からの感情推定の結果を操作し選択肢と一致させることで主人公の感情を選択する。
- 読み上げと表情の要素を同時に行い演技の入力とし、選択の分岐を行う。

これらの設計指針に基づき、提案手法ではウェブカメラ [9]を用いてプレイヤーの表情と発話を取得することで演技の入力を行う。プレイヤーはウェブカメラに向かって演技を行い、アドベンチャーゲーム内の選択肢を選択する。

3.2 実験用アドベンチャーゲームプログラム

実験用のアドベンチャーゲームをUnity (バージョン2020.3.18f1) [10]を用いて実装した。ゲームの趣向による個人差に配慮し、通常アドベンチャーゲームに含まれるキャラクターや背景の絵、音楽などを排除し、シナリオテキストのみを表示する形にした。ゲームの基本システムは通常のアドベンチャーゲームと同様である。プレイヤーはテキストを読み、マウスクリックでテキスト送りを行う。

実験で用いるシナリオは計3本、それぞれおよそ5分程度で終わる長さのものを用意した。内容に関しては、個人の趣向が反映されないよう、特殊性のない、議論の会話を中心としたものとした。具体的には、どのシナリオもプレイヤーキャラクターは大学新生であり、新入生歓迎会での

アイスブレイクとして周りの新入生とともにコンセンサスゲームを行う設定となっている。今回シナリオに用いた題目は月面での不時着、砂漠での不時着、雪山での遭難の3種である。実験中は題目内容の書かれたウィンドウをゲームウィンドウの横に配置し、常に見返せる状態になっている(図1)。

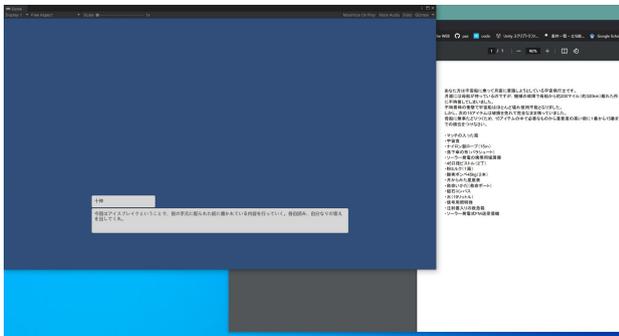


図1 実験中の画面構成。左側に実験用ゲームウィンドウ、右側にシナリオを補足するテキストが書かれたウィンドウを表示している。

音声認識では、ウェブカメラのマイクを用いて取得した音声データを、Unityにて文字列データに変換する。ここで返された文字列および、選択肢の文章を Bag of Words 形式に変換する。変換された発話内容と選択肢のコサイン類似度を求め、発話内容と最も一致する選択肢を求める。また、誤認識の可能性を最小限に抑えるため、シナリオ上での選択肢の文章は同じ単語は使用しないようにした。

表情からの感情推定には Python の paz ライブラリ [11] に含まれる Emotion classifier を使い、ウェブカメラにて撮影されたプレイヤーの表情の映像をリアルタイムで処理する。標準で推定される感情のうち、安定して意図した出力ができること、発話との両立が可能であることの2点を満たす感情は happy と sad のみであったため、この2感情で実装した。

また、Emotion classifier にはカメラ画像とともに一番可能性の高い感情を表示するインタフェースが出力される。このインタフェースは練習中のみ表示し、本番では非表示にした。

また、このプログラムにて出力された値は Unity に送られ、選択肢の選択に用いられる。Unity 側では、プレイヤーが提案手法の入力をしている間の happy と sad の値の累積により選択肢の選択を行う。

3.3 インタクションデザイン

提案手法のゲームは、選択肢が出るまでは通常のアドベンチャーゲームと同様である。選択肢では、最初にスペースを押すことで入力開始となる。すると、図2のように“now input”の文字列に加えて現在の感情を示す文字列が

表示される。ここで、選択肢の文章を読み上げながら意図した感情になるよう表情を作成する。その後、もう一度スペースを押すことで入力終了となる。音声認識および表情からの感情推定に用いられるのは入力開始から終了までの間である。選択完了後は再び通常のアドベンチャーゲームに戻る。

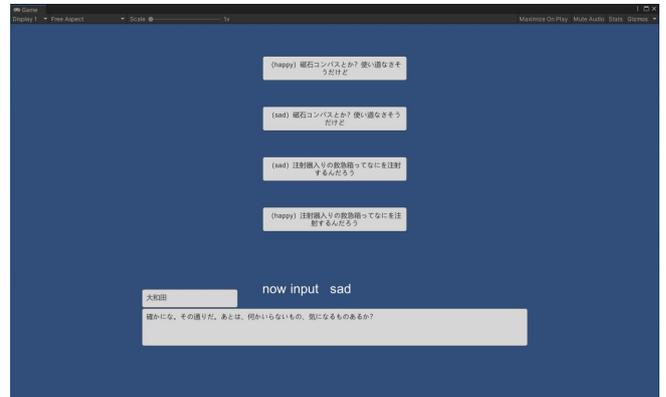


図2 実験用ゲームにおける演技入力中の画面。選択肢表示時に加え、“now input”と現在の感情を示す文字列が表示されている。

4. 実験

本研究において、提案手法によるユーザ体験効果を明らかにするために、実装したシステムを用いた実験室実験を行った。本章では、実験にあたって設定した仮説およびその仮説を検証するために実施した実験の内容についての説明を行う。

4.1 仮説

評価に当たっては、提案手法の効果として以下の2点の仮説を設定した。

- H1. 提案手法は従来手法と比べ没入感およびゲームに対する印象が向上する。
- H2. 提案手法は従来手法と比べ操作難度が高く、プレイヤーの負担になる。

提案手法である演技による入力手法では表情からの感情推定を用い、プレイヤーに場面に合った表情をとらせる。これにより2.1節で述べた表情フィードバック仮説の効果通り、シナリオ内で生まれる情動をさらに強化し、その結果ゲームに対する没入感およびゲームに対する印象の向上につながる。しかし、従来手法のマウスクリックやボタン選択と比べると入力に必要な技術および労力が大きく、それだけプレイヤーの負担になると考えた。

4.2 実験条件

入力手法の条件はマウス、読み上げ、演技の3通りである。マウス条件は選択肢の選択をマウスクリックで行うも

の、読み上げ条件は選択肢の選択を発話のみで行うもの、演技条件は選択肢の選択を発話および表情からの感情推定の両方で行うものである。入力手法に読み上げ条件を加えた理由は、本研究では表情フィードバックを基にした表情の作成を入力手法とした場合の効果を明らかにするためである。演技条件は発話+表情の作成であるが、マウス条件と演技条件の比較だけでは、効果が出た場合も、それが発話+表情の作成によるものなのか、発話だけによるものなのか判断できない。そのため発話のみの読み上げ条件を加えた。

シナリオ条件は、3章で述べた通りの月、砂漠、雪山の3通りである。各参加者には3回のタスクで、すべての入力手法およびすべてのシナリオを体験してもらう。タスクを行う順番は順序効果に配慮し、入力条件に関しては全通り、シナリオ条件に関してはラテン方格に基づいた順序の組み合わせで行った。

4.3 実験タスク

最初に、この実験の目的および好きな時に休憩をしてよいことを説明した後、表情の訓練を行った。これは、pazライブラリのインタフェースを見せながら、実験参加者にhappyとsadを安定して出せるようになってもらう訓練である。また、片方しか出せない実験参加者には、出せる方の感情の数値を一部出せない方の感情の数値に変換することで補正をかけ、両方の感情を出せるように調整した。

その後、入力手法の練習、ゲームプレイ、アンケート回答を1セットとし、これを3セット行った。入力手法の練習は実験参加者が入力手法の理解ができるまで何度も行う。また、演技条件の入力練習時にはpazライブラリのインタフェースを実験参加者に見せながら行う。実験用ゲームは1人用を想定して作成したため、実際のシナリオを用いたゲームプレイ中は実験担当者は部屋の外に出て、実験参加者が1人になる環境にした。また、各シナリオに関して、選択肢は3回出てくる。すべて終了した後に半構造化インタビューを行う。

4.4 実験参加者

実験参加者は計21名(M=19, F=2)、21-27歳(mean=23.238, sd=1.540)の大学生および大学院生である。ただし、最初に実験を行った3名に関して実験環境に問題があったため、分析から外し18名で分析を行った。

4.5 評価項目

本実験において、仮説を検証するために3つの評価項目を設定した。1, 2個目はH1を、3個目はH2を検証するための項目である。

- ゲームに対する没入感
- ゲームに対する印象

- ゲームの入力に対する緊張

ゲームに対する印象は、面白い等ポジティブな評価を印象が高い、つまらない等ネガティブな評価を印象が低いとする。評価にあたっては、ゲームに対する没入感、印象、緊張の評価を行うために2種類のアンケートを実施した。また、全体を通し意見を求めるために半構造化インタビューを行った。

4.5.1 IMI アンケート

IMI (Intrinsic Motivation Inventory) [12]とは、内発的動機づけの考えに基づいて作成された、7段階のリッカート尺度で評価するアンケートである。このアンケートにおいて、不要な質問項目を削除した先行研究[13]の内容を和訳し、必要な評価項目の質問を残したものを本実験でアンケートとして用いた。本実験ではinterest-enjoyment, tension-pressureの2つの指標を用いる。interest-enjoymentはゲームに対する印象、tension-pressureはゲームの入力に対する緊張を評価する。

4.5.2 GEQ アンケート

GEQ (Game Experience Questionnaire) [14]とは、ゲームを評価する際に広く使用されている、5段階のリッカート尺度で評価するアンケートである。本実験ではcore moduleを和訳したもののから、必要な評価項目の質問を残したものをアンケートとして用いた。本実験ではimmersive, flow, positive, negativeの4つの指標を用いる。immersive, flowはゲームに関する没入感を、positive, negativeはゲームに対する印象を評価する。

4.5.3 半構造化インタビュー

半構造化インタビューでは、入力手法の好みとその理由、ゲーム経験、入力難度および演技に対する考え方を質問した。

5. 結果および考察

本章では、実験の各評価項目についての結果を示し、考察を行う。また、アンケートの評価に対する検定において、p値が0.05未満であることを統計的に有意とみなした。

5.1 IMI アンケート

5.1.1 Interest-Enjoyment に関する結果および考察

アンケート結果のデータを箱ひげ図にしたものを図3に示す。シャピロウィルク検定の結果、データに正規性が認められなかった($p < 0.05$)。そのため、ノンパラメトリック検定を行った。クラスカル・ウォリス検定を行い、その結果、すべての条件間に有意差は見られなかった。すなわち、入力手法を変化させてもゲームに対する興味、および面白さの印象に有意差は見られなかった。原因として、シナリオの淡白さにあると考える。シナリオ内容は4.1.1項で述べた通り、プレイヤー個人の趣向が評価に反映されないよう、シナリオは議論の会話によって構成された、感情の

変動が小さいものとした。しかし、その結果プレイヤーの感情は入力手法を変えても変化せず、ゲームに対する印象に差が見られなかったと考えられる。

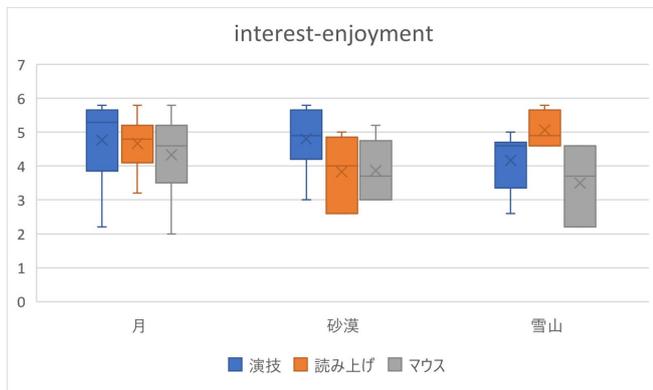


図 3 interest-enjoyment アンケート結果の箱ひげ図。×はデータの平均値を示す。

5.1.2 Tension-Pressure に関する結果および考察

アンケート結果のデータの平均を棒グラフにしたものを図 4 に示す。シャピロウィルク検定の結果、データに正規性が認められた ($p > 0.05$)。また、バートレット検定の結果データに等分散性が認められた ($p > 0.05$)。そのため、二元配置分散分析を行った。その結果、入力手法および入力手法とシナリオの交互作用が有意であった (入力手法: $p < 0.001$, 交互作用: $p < 0.05$)。交互作用で有意差が見られたため、単純主効果の検定を行った。その結果、シナリオが砂漠の際の入力手法 ($p < 0.05$)、雪山の際の入力手法 ($p < 0.01$)、および入力手法が読み上げの際のシナリオ ($p < 0.01$) に有意差が見られた。そのため、有意差が見られた条件で多重検定を行った結果、シナリオが砂漠の際の入力手法ではマウス-演技条件間 ($p < 0.05$)、シナリオが雪山の際の入力手法ではマウス-演技条件間 ($p < 0.05$) およびマウス-読み上げ条件間 ($p < 0.01$)、入力手法が読み上げの際のシナリオでは雪山-砂漠 ($p < 0.01$) および雪山-月 ($p < 0.05$) で有意差が見られた。

シナリオが砂漠の際の入力手法、および雪山の際の入力手法で有意差が見られた組み合わせでは、マウス条件よりも演技条件もしくは読み上げ条件の緊張の値が有意に高く、入力時の緊張が大きいと言える。しかし、この結果が見られたのは一部の条件間のみである。また、入力手法が読み上げの際のみ雪山の値が有意に大きい理由として、シナリオ上の選択肢の文章の内容が考えられる。各シナリオで選択肢は 3 回設けられ、すべて選択肢の文章の意図はシナリオ間で差のないように作られている。しかし、3 回目の選択肢でのみ、月、砂漠シナリオと比べ雪山シナリオの文章が 1 単語分長く設定されていた。演技条件では表情にも気を配る必要があるため文章の長さの差は結果として見られなかったが、読み上げ条件では文章の長さが入力時の

緊張に繋がり、結果に表れたのだと考える。

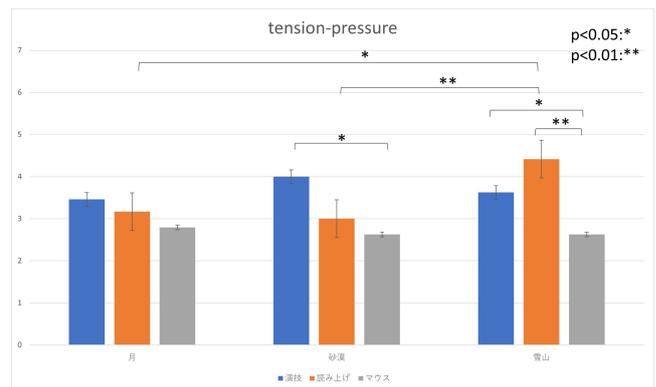


図 4 tension-pressure アンケート結果。エラーバーはデータの標準誤差を示す。また、有意差が出た条件の間に*を示す。

5.2 GEQ アンケート

5.2.1 Immersive に関する結果および考察

アンケート結果のデータの平均を棒グラフにしたものを図 5 に示す。シャピロウィルク検定の結果、データに正規性が認められた ($p > 0.05$)。また、バートレット検定の結果データに等分散性が認められた ($p > 0.05$)。そのため、二元配置分散分析を行った。その結果、入力手法の主効果が有意であり ($p < 0.05$)、シナリオの主効果および入力手法とシナリオの交互作用は有意でなかった。入力手法に関して有意差が見られたため、チューキー HSD 検定で多重比較を行った。その結果、マウス条件と演技条件の間に有意差が見られた ($p < 0.05$)。

結果として、マウス条件に比べ、演技条件の方がゲームに対する没入感が高いことが分かった。しかし、読み上げ条件と演技条件の間にはゲームに対する没入感の差は見られなかった。読み上げ条件と演技条件の間に差が見られなかった理由としては、2.1 節で述べた表情フィードバックの効果を十分に発揮できないシナリオにあると思われる。実験で用いたシナリオは感情の変化には不十分であり、表情フィードバックによる情動の強化が起こらず、発話に表情の作成を追加した効果が見られなかったと考える。

5.2.2 Flow に関する結果および考察

アンケート結果のデータの平均を棒グラフにしたものを図 6 に示す。シャピロウィルク検定の結果、データに正規性が認められた ($p > 0.05$)。また、バートレット検定の結果データに等分散性が認められた ($p > 0.05$)。そのため、二元配置分散分析を行った。その結果、入力手法の主効果が有意であり ($p < 0.01$)、シナリオの主効果および入力手法とシナリオの交互作用は有意でなかった。入力手法に関して有意差が見られたため、チューキー HSD 検定で多重比較を行った。その結果、マウス条件と演技条件の間に有意差が見られた ($p < 0.01$)。

結果として、マウス条件に比べ、演技条件の方がゲーム

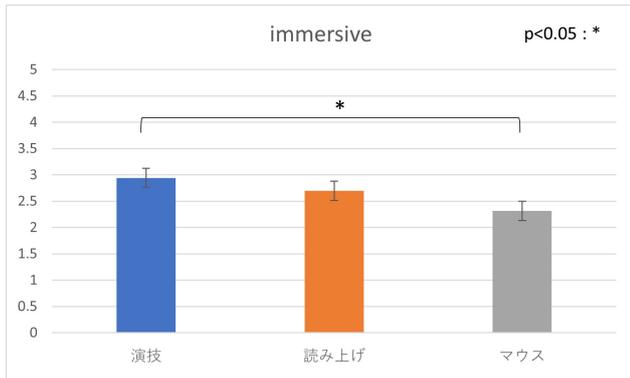


図 5 immersive アンケート結果. エラーバーはデータの標準誤差を示す. また, 有意差が出た条件の間に*を示す.

に対する没入感が高いことが分かった. しかし, 読み上げ条件と演技条件の間にはゲームに対する没入感の差は見られなかった. 読み上げ条件と演技条件の間に差が見られなかった理由としては, 5.4.2 項の Immersive に関する結果および考察と同様に, 感情の変化が不十分なシナリオにより表情フィードバックによる情動の強化が起らなかったためであると考え.

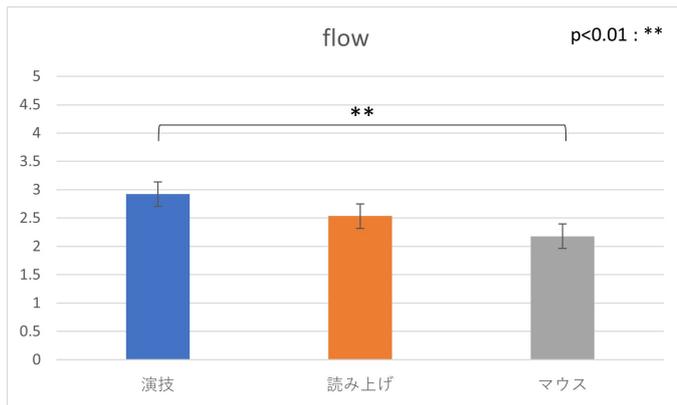


図 6 flow アンケート結果. エラーバーはデータの標準誤差を示す. また, 有意差が出た条件の間に*を示す.

5.2.3 Positive に関する結果および考察

アンケート結果のデータを箱ひげ図にしたものを図 7 に示す. シャピロウィルク検定の結果, データに正規性が認められなかった ($p < 0.05$). そのため, ノンパラメトリック検定を行った. クラスカル・ウォリス検定を行い, その結果, すべての条件間に有意差は見られなかった.

この結果, 入力手法を変化させてもゲームに対するポジティブな印象の評価に有意差が見られないことが分かった. 原因としては, 5.4.1 項の Interest-Enjoyment に関する結果および考察と同様に, 本実験のシナリオが淡白であったため, 入力手法の変化がゲームへの印象の変化に繋がらなかったと考える.

5.2.4 Negative に関する結果および考察

アンケート結果のデータを箱ひげ図にしたものを図 8 に

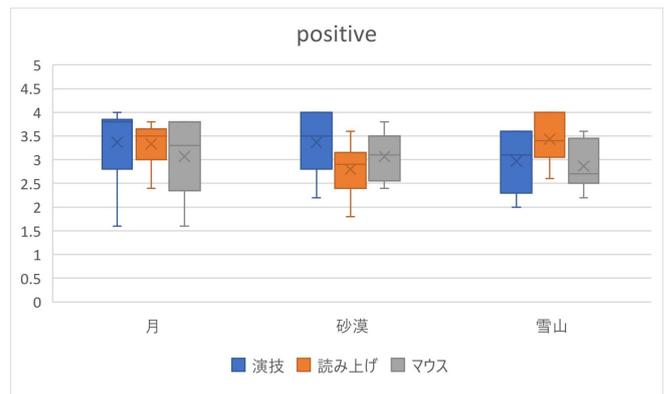


図 7 positive アンケート結果の箱ひげ図. ×はデータの平均値を示す.

示す. シャピロウィルク検定の結果, データに正規性が認められなかった ($p < 0.05$). そのため, ノンパラメトリック検定を行った. クラスカル・ウォリス検定を行い, その結果, すべての条件間に有意差は見られなかった.

この結果, 入力手法を変化させてもゲームに対するネガティブな印象の評価に有意差が見られないことが分かった. 原因としては, 5.4.1 項の Interest-Enjoyment に関する結果および考察と同様に, 本実験のシナリオが淡白であったため, 入力手法の変化がゲームへの印象の変化に繋がらなかったと考える.

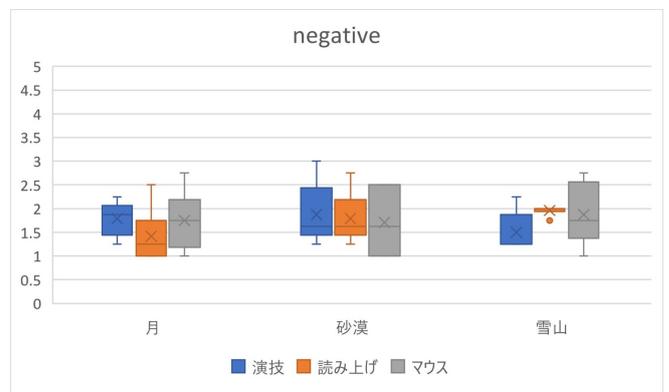


図 8 negative アンケート結果の箱ひげ図. ×はデータの平均値を示す.

5.3 半構造化インタビュー

主観による好みの入力手法に対する質問の結果は, マウス条件と答えた人は 7 人, 読み上げ条件は 7 人, 演技条件は 4 人であった. マウス条件が一番好みであると答えた人の理由としては, 入力が容易である, 反応がはやい, 慣れている, 入力の確実さに対する安心感等が挙げられた. 読み上げ条件が一番好みであると答えた人からの理由としては, 演技条件では表情作成が難しく読み上げ条件の方がゲームに集中できた, マウスよりもゲームへの参加を感じた, 操作がそこまで苦でない等が挙げられた. 演技条件が

一番好みであると答えた人からの理由としては、マウス入力は容易だが面白味がない、感情が反映されるのは新しく面白い等が挙げられた。

また、主観評価と過去のゲーム経験の関連性についてはほとんど見られなかった。過去にアドベンチャーゲームをプレイしたことがある人は、慣れているためマウス条件を好む人もいれば、新しさから読み上げ条件、演技条件を好む人もいた。普段からゲームをしない人も、入力の容易さからマウス条件を好む人もいれば、体験の豊かさから読み上げ条件、演技条件を好む人もいた。

演技入力の精度について、1本のシナリオ中3回ある選択肢の選択において平均0.5回の入力ミスがあった。すべてが意図した感情の入力の失敗であり、表情の作成に関して安定した入力が困難であると考えられる。

演技に対する考え方と本手法の演技に対する評価に関連性は見られなかった。演技に対してネガティブな考えを持つ人でも否定的な意見ばかりではなく、“ほかの人が近くにいると恥ずかしかっただろうが、ゲーム相手だと楽しかった”と肯定的な意見を持つ人もいた。一方で演技に対してポジティブな考えを持つ人でも肯定的な意見ばかりではなく、労力や入力精度への不安からの否定的な意見を持つ人もいた。

最後に、ゲーム、システム面への改善案として多かったものを述べる。音声入力の反応の遅さが目立つ。演技条件での表情作成は、演技というよりは表情のトレーニングであり、不自然なほど過度に表現しなければならない。コンセンサスゲームという題材はプレイヤー自身の意見をもつ原因になり、選択肢の文章に対して自分の意見と違うという不満を起こす可能性がある。声に対する感情推定技術を用いて、発話のみで感情の入力までできると、表情の作成をする労力がなくなり良い。以上のような意見が得られた。

5.4 実験結果に対する考察のまとめ

H1に関して、マウス条件と比較し、演技条件のゲームに対する没入感の向上は見られた。しかし、読み上げ条件と演技条件の間には差は見られなかった。また、ゲーム自体への印象の向上は見られず、ユーザ体験がよくなったとは言いきれない。そのため、H1の仮説は部分的に支持されたと考える。

H2に関して、マウス条件の方が読み上げ、演技条件より緊張の値が小さく、読み上げ、演技の入力時にマウスと比べ緊張が生まれている。一方、シナリオによる差も見られた。また、読み上げと演技の間には有意差が見られなかった。結論として、H2の仮説は部分的に支持されたと考える。

また、半構造化インタビューを行い、実験参加者の主観的意見を集めた結果、アンケート調査の結果とは異なり、マウス条件と読み上げ条件を好む人が多かった。要因とし

ては、演技入力の際の表情作成の労力が大きく、過度な表現を求められていたことが考えられる。また、演技入力の失敗数から提案手法の入力難度が高いと考えられる。

6. 本研究の制約、今後の課題および展望

本章では、本提案手法に対する議論を行い、本研究の制約、今後の課題および展望を述べる。本研究で用いた表情からの感情推定を行うライブラリでは、口、目、眉がカメラに映らないと感情推定を行うことができない。また、意図した感情を入力することが難しい。標準で識別される7感情のうち、安定して入力できる感情は2-3つにとどまる。また、実験をして感じた点として、表情には個人差が大きく、その人に合った数値の調整が必要であると感じた。

また、安定して入力できるものの、安定させるためには過度に表情を作らなければならない。その表情を保ち続けるのは労力が大きく、否定的な意見が多く見られた。また、演技という観点でも、現在要求されている表情は大袈裟であり、不自然であるという意見も見られた。

現在の表情からの感情推定では、目、眉とともに口の形状が用いられている。これは提案手法の、発話と同時に表情を作成することに対して致命的である。

この問題の解決策として、口の形状を用いない表情からの感情推定を行う、もしくは表情ではなく音声から感情推定を行うことが考えられる。前者に関しては、口の形状を用いないことにより、発話内容によって推定結果が左右されることはなくなる。しかし、感情推定の精度は本実装で用いたものよりも低くなるのが考えられ、プレイヤーが意図した感情の入力が行えるかは不明である。また後者に関しては、入力時に表情を作成する必要がなくなる。そのため、プレイヤーは問題なく意図した感情の入力が行えると考えられる。しかしその場合、2.1節で述べた表情フィードバック仮説は用いられず、本実験で明らかとなった、演技入力による没入感の向上が適用されるかは不明である。

今回実験で用いたシナリオは、議論の会話が中心の、淡白で短く、感情の変動には不十分であった。シナリオの流れも、選択肢の後分岐するのは1, 2パラグラフで、選択肢に対してNPCが正しい反応を返した後すぐにすべての分岐が統一されてしまうため、選択によるシナリオの変化が通常アドベンチャーゲームに用いられるシナリオと比べ小さい。この部分を改善し、より感情の動くシナリオにすることで、表情フィードバックによる情動の強化が効果的に適用され、今回変化がなかったゲームに対する印象の指標も変化があると考えられる。また、アドベンチャーゲームの中でも、恋愛シミュレーション等、演技と相性のよいジャンルは存在すると考えられる。

半構造化インタビューにて、一般的には演技はネガティブなとらえ方をされやすいという意見があった。実際に演技を入力とするゲームを作成した場合、演技をネガティブ

にとらえる層には一切興味を持たれない、ターゲットを非常に絞ったゲームになる。実際に製品とすることを考えた際には非常に大きな制約となる。ただし、本研究における実験参加者は演技をポジティブにとらえている人の方が多数派であり、演技をネガティブにとらえる人も、理由として人に見られるのが恥ずかしく、他人のいない環境でゲームを相手にする場合は特に気にならない人が多かった。実際に、演技には苦手意識を持っているが、このゲームの演技入力には面白かったという意見も見られた。

本研究の実験では、実装に用いた表情からの感情推定の精度および音声認識の性能が十分でなかった。そのため、過度な表情の作成が求められ、入力完了から選択が反映されるまでの時間が長い問題点が明らかになった。今後は実装における表情からの感情推定の精度および音声認識の性能を向上させ、この課題を解消する。

また、本実装で用いた表情からの感情推定は口の形状を要素として含むことから、表情の作成と発話を同時に行うことが困難であることが明らかになった。今後、表情からの感情推定にて口の形状を用いないものを作成することでこの課題を解決する。または、表情からではなく発話の音声から感情推定を行う手法を実装することでこの課題を解決する。

また、本研究の実験で用いたシナリオは淡白で短く、感情の変動に不十分であった。そのため、表情フィードバックを用いた情動の強化につながらず、提案手法がゲームに対する印象にもたらす影響を検証するのに適していなかったと考える。今後は入力手法の違いによる印象の変化を明らかにするため、シナリオ内容を実際のアドベンチャーゲームに用いられるものと同様の、感情の起伏に富んだものにして再び実験を行う。また、演技入力と相性のよいシナリオのジャンルを明らかにするため、演技入力で複数ジャンルのアドベンチャーゲームをプレイしてもらった実験を行う。

7. おわりに

アドベンチャーゲームにおいて表情からの感情推定技術を用い、発話と表情の作成を組み合わせた演技を入力手法とするシステムを提案した。提案手法を用いた実験により、以下の結果が得られた。

- 従来手法と比較し、ゲームに対する没入感が向上した。
- 従来手法と比較し、ゲームに対する印象の変化は見られなかった。
- 従来手法と比較し、操作難度が高く、入力においてプレイヤーの負担となる。

また、本研究の実験で用いたシステムには、音声認識の処理速度、表情からの感情推定の精度の低さからくる過度な表情の作成を要する点および発話との両立が困難である点、シナリオ内容などの課題がある。そのため、今後は音

声認識および表情からの感情推定プログラムの再検討、シナリオ内容の改善、提案手法およびアドベンチャーゲームシステムの再検討を行っていく予定である。

参考文献

- [1] Lazarus, R. S. and Lazarus, B. N.: *Passion and reason: Making sense of our emotions*, Oxford University Press, USA (1994).
- [2] Cornelius, R. R.: *The science of emotion: Research and tradition in the psychology of emotions.*, Prentice-Hall, Inc (1996).
- [3] James, W.: What is an emotion? *Mind*, os-IX, 188-205 (1884).
- [4] Tomkins, S. S.: *Affect Imagery Consciousness*, vol. 1, The Positive Affects (London: Tavistock, 1962); and Paul Ekman, *Emotion in the Human Face*.
- [5] Strack, F., Martin, L. L. and Stepper, S.: Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis., *Journal of personality and social psychology*, Vol. 54, No. 5, p. 768 (1988).
- [6] Tsujita, H. and Rekimoto, J.: Smiling Makes Us Happier: Enhancing Positive Mood and Communication with Smile-Encouraging Digital Appliances, *Proceedings of the 13th International Conference on Ubiquitous Computing, UbiComp '11*, New York, NY, USA, Association for Computing Machinery, p. 1-10 (online), DOI: 10.1145/2030112.2030114 (2011).
- [7] Lankes, M., Riegler, S., Weiss, A., Mirlacher, T., Pirker, M. and Tscheligi, M.: Facial Expressions as Game Input with Different Emotional Feedback Conditions, *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, ACE '08*, New York, NY, USA, Association for Computing Machinery, pp. 253-256 (online), DOI: 10.1145/1501750.1501809 (2008).
- [8] Robinson, R. B., Reid, E., Fey, J. C., Depping, A. E., Isbister, K. and Mandryk, R. L.: *Designing and Evaluating 'In the Same Boat', A Game of Embodied Synchronization for Enhancing Social Play*, p. 1-14 (online), available from (<https://doi.org/10.1145/3313831.3376433>), Association for Computing Machinery (2020).
- [9] Logicool: Logicool HD Pro Webcam C910, <https://www.logicool.co.jp/>.
- [10] Technologies, U.: Unity, <https://unity.com/>.
- [11] Arriaga, O., Valdenegro-Toro, M., Muthuraja, M., Devaramani, S. and Kirchner, F.: Perception for Autonomous Systems (PAZ) (2020).
- [12] Ryan, R. M.: Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory., *Journal of personality and social psychology*, Vol. 43, No. 3, p. 450 (1982).
- [13] McAuley, E., Duncan, T. and Tammen, V. V.: Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis, *Research Quarterly for Exercise and Sport*, Vol. 60, No. 1, pp. 48-58 (online), DOI: 10.1080/02701367.1989.10607413 (1989). PMID: 2489825.
- [14] IJsselstein, W. A., De Kort, Y. A. and Poels, K.: The game experience questionnaire (2013).