

声の高低を用いた和音学習ゲームの開発

福嶋穂倅^{†1} 川合康央^{†1}

概要: 本研究では、ユーザの声の高低を利用した和音学習ゲームの開発を行ったものである。和音は、音楽の学習や演奏する上で、必要となる知識であるが、その取得は難しい。そのため、より効率的に学習する必要があると考え、音程解析を利用した学習形式の RPG ゲームの開発を行うこととした。音程解析をゲーム操作に組み込むことによって、より直感的な学習が可能となる。

1. はじめに

1.1 概要

本研究では、声の高低を用いた操作が可能なゲームコンテンツの開発を行った。既存のゲーム操作では、ボタンやジョイスティックなどの入力デバイスを使用している。一方、先行研究では、日常動作である歩行を元にしたゲームインタフェースや[1]、表情をパラメータにしたゲームインタフェース[2]など、様々な身体情報を用いたゲームインタフェースの提案が行われている。これらの研究を基に、声による直感的な操作が可能なゲームインタフェースについて検討し、新たな操作方法として提案する。

また、我々は通常、自分自身の身体を日常的に動かしていることに比して、ゲーム内の仮想の身体が、必ずしもユーザの感覚と一致しているものではない。ゲーム内の仮想的な身体を、ユーザが思うように動かさなければ、操作やその上達も難しく、結果として、ゲームからの離脱を招く可能性がある [3]。そのため、プレイヤーの実際の感覚がゲーム内の感覚と直感的に結びつくように、インタフェースを工夫をしなければならないと考える。

さらに、プレイヤーがゲームを面白いと感じる要素として、初期の頃には、ユーザのスキルが低くても超えられるハードルがあること、技能や知識が蓄積し高まっていくと感じられること、出来たという手ごたえを得られること等が挙げられており[4]、これらの研究結果からも、ゲームの操作性はそのゲームに対して大きな影響を与えるものとして考えられる。これらの先行研究を基に、本研究では、現実での感覚を、直感的にゲームに反映できる音声情報を利用したゲームデザインの開発を行った。

2. 開発手法

2.1 開発環境

本研究では以下の開発環境を使用して開発を行った。

表 1 開発環境

プログラミング言語	C#
開発ツール	Unity 2021.3.9.fl Visual Studio Code
ハードウェア	Windows10 Pro
グラフィックツール	Clip Studio Paint Photoshop Maya
その他ツール	AudioPitchEstimatorFor Unity

本研究では、開発環境として、ゲームエンジン Unity を用いることとした。Unity は、オーディオファイルに対して高速フーリエ変換を行う関数が存在している。また、Unity 内で使用可能な音声処理システムとして、本研究では AudioPitchEstimatorForUnity を使用することとした。この AudioPitchEstimatorForUnity を使用することによって、音声の音程を分析することが可能である。また、開発言語として、C#を使用し、Visual Studio Code を用いてコーディングを行うこととした。また、キャラクタや背景等のデザイン制作には、ペイントソフトウェア Clip Studio Paint、3DCG ソフトウェア Maya を使用し、3D モデルのテクスチャ制作には AdobePhotoshop を使用した。

2.2 コンテンツ概要

音階を使用することから、ゲームデザインは音楽とかけ合わせることが効果的であると考えた。しかし、単にリズムゲームとするのではなく、今回提案するインタフェースに RPG (Role-Playing Game) をかけ合わせることによって、インタラクティブなコンテンツになるよう開発を行った。本研究で開発したコンテンツは、ゲーム操作を声の音階のみで行うものとした。

^{†1} 文教大学情報学部情報システム学科

またゲームは音楽におけるドからシ(C,D,E,F,G,A,B)までの音がパソコンに接続するキーボードの(C,D,E,F,G,A,B)に対応することに目をつけ、キーボードと音、例えばドの音であればCのキーを、レの音であればDのキーを使用することにより DAW 等で用いられる Qwerty キーボードよりもわかりやすく直感的に扱いやすいインタフェースになるのではと考え開発を行うこととした

PCに接続したマイクインタフェースから音声を取得し、Unity で音階の分析を行い、そこからゲーム操作に繋がる処理を行うこととした。

このコンテンツは RPG がメインのゲームである。RPG のバトルシーンにおいて主な操作である攻撃、防御、回復の 3 つの操作について声を用いて行うこととした。ゲームデザインを考え、各操作を行うためには、画面下部にある平行四辺形のゲージを満たす必要がある。

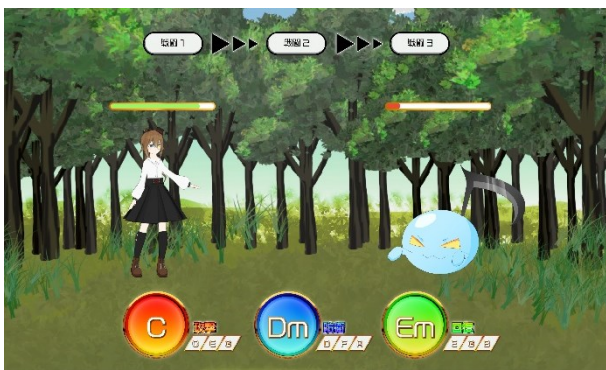


図 1 ゲーム画面

ゲームの動作画面を図 1 に示す。ゲーム上で攻撃を行う場合には、攻撃ボタンである C のコードを構成している音すべてを満たさなければならない。すなわちプレイヤーは、C コードの構成音である C, E, G の音を発生する必要がある。一方で、プレイヤーがすべての音を補助なしで発声することは困難だと考え、該当する音のアルファベット音名のキー、例えばドの音が知りたい場合にはアルファベット音名である C のキーをキーボードから入力することで、その音を鳴らすことができる機能を設けた。



図 2 UI について

プレイヤーがマイクから音声を入力すると、画面の上部に音程が図 2 のように表示される。図 2 の状態であれば、C の音を認識したため、攻撃ボタンのゲージである C, E, G のうちの C を満たしたこととなる。このように、すべてのゲージが満たされると、敵キャラクターに対してプレイヤー側からの攻撃が行われる。

先行研究から音声の認識精度について、認識率の上昇や誤認識の低減をはかる方法としてインストラクションが大事であるといった研究結果がある[5]。そこで、音声情報を可視化することで、プレイヤーが現在発している音と、操作に必要な音との差が分かるよう、画面上部に現在自分が発している音が、どの帯域のどの音なのかを示すようなインタフェースデザインとした。



図 3 画面要素

画面全体のデザインは、図 3 のようになる。画面上部には、マップの進捗具合が表示される。また、各キャラクターの上部には、体力である HP ゲージが表示され、敵キャラクターの HP ゲージが 0 になると次の戦闘へ進み、プレイヤーキャラクターのゲージが 0 になると、ゲームオーバーとなる。攻撃を受けると HP ゲージは減少していく。

2.3 システム概要

続いてシステムの概要と画面遷移について、システム概要図を図 4 に示す。

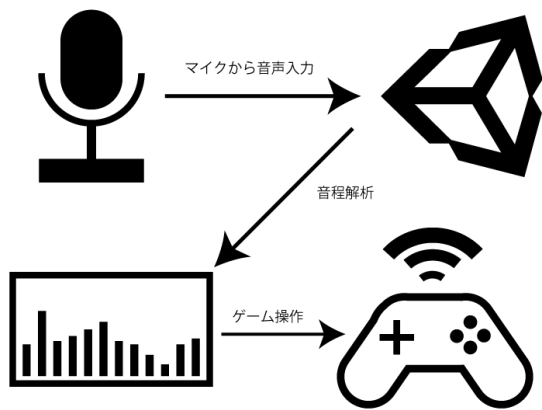


図4 システム概要図

PCに接続されたマイクから音声情報を取得し、これをゲームエンジンUnity上で処理を行う。この際に、AudioPitchEstimatorForUnityを使用して、取得した音声の音階を分析する。この分析結果を基に、ゲームシステムへの反映を行っている。

次にゲーム全体のフローチャートを図5に示す。

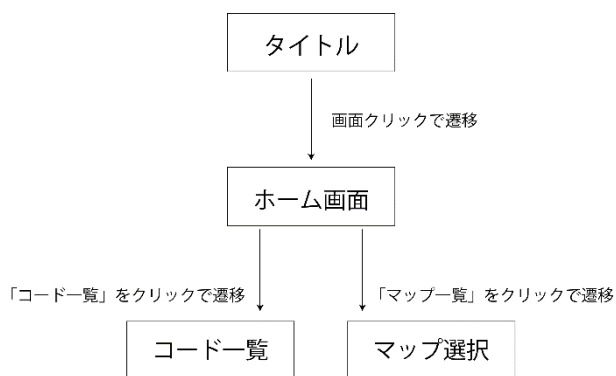


図5 ゲームフローチャート

タイトルから始まり、画面をクリックすることでゲームのホーム画面へと移動する。このホーム画面からは、ゲーム内で使用しているコードの確認が行えるコード一覧の画面と、戦闘を行うために使用するマップを選択するマップ選択画面の2つへと遷移することができる。

ゲームサイクルについては図6のようになっている。マップを選択するとゲームが開始される。続いて戦闘シーンへと移行する。戦闘は全3回のうち2回が通常の戦闘、1回がボス戦闘となっている。1回目の戦闘が終わると2回目の戦闘へと移行する。最後の戦闘を終えるとゲームクリアの判定になり、次のマップが解放されるというのが一連のゲームサイクルとなっている。

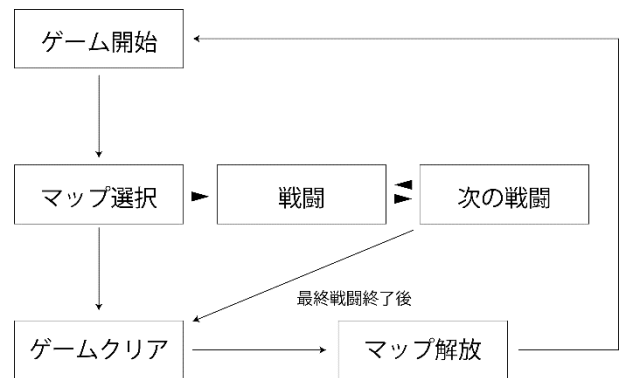


図6 ゲームサイクル

3. 課題と考察

3.1 考察

今回の研究では、ゲーム操作に音声を用いるコンテンツの制作を行った。RPGに音声インタフェースを組み合わせることによって、プレイヤーが面白いと感じる要素の1つである技能や知識の蓄積や、クリア後に何度も楽しめるといった、ゲームインタフェースに必要ないくつかの要素を満たすことができた[4]。現実での我々自身の感覚を、ゲーム内に反映させることによって、よりプレイヤーに対して面白いと感じられるゲームインタフェースの開発を行うことが可能である。今回提案したような、ゲームパッドやキーボードなどの物理的な入力デバイスを用いないゲーム操作に関しては、これまでも先行研究が存在している。今後もこのような直感的に操作が行えるインタフェースは、ゲーム操作だけでなく様々な形で検証され実用化されていくと考えられる。

3.2 課題

一方でいくつかの課題もある。先行研究の1つである表情をゲームのパラメータとして応用した例では、「幸福」の表情が最も操作しやすく、「悲しみ」「恐怖」「怒り」「嫌悪」の表情で操作することが困難であったとしている。その理由として、これらの表情を意図的に表現しているわけではなく、理性的に隠し抑えている働きがあるためだとしている[2]。本研究でインタフェースとして用いた声に関しても、普段から自分が発声している音域以外は出す機会が少ない。そのため、同じように自分が普段使用している音域以外では操作がしにくいという課題が残った。今後は声による操作について、ユーザ評価を行い、より直感的な操作を考える必要がある。ゲームをプレイする前に声のキャリブレーションを行い、自分が発声できる範囲内で操作できるインタフェースや、音階以外にも音の強弱で操作するインタフェースなどの展開が考えられる。

4.まとめ

今回の研究では、ゲームエンジン Unity と、音程の検出を行うために AudioPitchEstimatorForUnity を使用し、マイクから取得した音声を元に音階を分析し、それによって操作を行うゲームコンテンツの開発を行った。このようにゲーム操作に非言語音声を用いることで、ユーザに新たな体験を提供することができると考えられる。

一方で、多くのユーザが扱うインタフェースとしてはいくつかの改善点があるため、引き続き開発を続けて行く。今回は音声と RPG を組み合わせたゲームコンテンツであったため、使用する音の高さについては絶対の音程を利用した。しかし、様々な人が使えるインタフェースとして使用する場合には、個人の音声を最初に検出し、そこからの相対の音程を利用することが望ましいと考えられる。

また、本研究では非言語音声を取り扱ってきたが、これらを自然言語処理と組み合わせることによって、さらにゲーム操作におけるインタフェースの幅は広がり、より魅力的な体験を得ることができると考えられる。

参考文献

- [1] 石井優衣, 栗原一貴, 湯村翼: 日常動作を複数の操作に再利用する「ながらゲーム」の提案, 情報処理学会インタラクティブセッション 2017, pp.325-327 (2017).
- [2] 丹羽将康, 渡邊 恵太: Facialaction: 表情をパラメータにしたインタラクションの提案とその応用, エンタテインメントコンピューティングシンポジウム 2020 論文集, pp.115-117 (2020).
- [3] 築瀬洋平, 鳴海 拓志: 誰でも神プレイできるジャンプアクションゲーム, 日本バーチャルリアリティ学会論文誌, Vol.21, No.3, pp.415-422 (2016).
- [4] 仲村真廣, 横田直明, 遠藤雅伸: プレイヤーがゲームを面白いと感じる要素に関する定性調査分析, 日本デジタルゲーム学会 2018 年度夏季研究発表大会予稿集, pp.91-92 (2018).
- [5] 相原政徳 小田嶋 和幸 畑岡 信夫: カーナビにおける音声インタフェースの実車評価, FIT2009 (第 8 回情報科学技術フ