

コトダマボウリング：音声認識と加速度センサーを用いたインタラクティブクイズゲーム

白田 充^{1,a)} 竹内邦宇¹ 馬場章 相澤 清晴¹

概要： 加速度センサーと音声認識を用いることで新たなインタラクションの形状をもったゲームを提案する。クイズゲームの解答を入力する箇所に音声認識を利用することで従来の入力方式であるボタンを押す操作やタッチスクリーンでの操作より、煩雑な操作をなくし円滑に解答できる仕様にした。また、操作デバイスの形状をボウリングのボール形状にすることでゲーム内容にも適合したボールを振るという違和感のない行動が行われ、自然で現実的な体験を齎すだろうと考えられる。本作品を展示した3つの場所での体験者の意見を通じて考察した。

Kotodama Bowling : Interactive Quiz Game with Speech Recognition and Accelerometer

MITSURU HAKUTA^{1,a)} KUNITAKA TAKEUCHI¹ AKIRA BABA KIYOHARU AIZAWA¹

Abstract: We propose a novel video game with the interaction by using speech recognition and accelerometer. By using speech recognition at where users input the answer of the quiz game, we made a system that eliminate complicated operations and can answer smoothly than by pressing buttons that are traditional input methods or touch screen control. In addition, by making the shape of the control device into a ball shape of bowling, behaviors that do not feel strange to the ball that also matches the contents of the game are performed, and it is thought that it will bring a natural and realistic experience. This work was exhibited at the three places and we considered through the opinion of participants.

1. はじめに

近年、ビデオゲームは多種多様な操作形態が存在している。数多くのビデオゲームが配信されているスマートフォンではタッチスクリーンを介した操作が行われるが、家庭用ゲーム機では従来のコントローラーを介した操作が多く行われる。また、パーソナルコンピュータでは主にキーボードやマウスを介して操作を行い、アミューズメントパークでは特殊な構造をした操作デバイスを介して操作を行うビデオゲームも存在する。

上記のように操作形態が多く存在する中、タッチスクリーンを介した操作を行うスマートフォンのビデオゲー

ム市場は急激に成長している。この市場成長に関してはスマートフォンの普及率やビデオゲーム自体を購入する必要がなく、無料で遊ぶことが可能な点、店頭で購入するといった外出してその場所まで行動に出る必要がなく、オンライン環境であれば即座にダウンロードして遊ぶことが可能な点などが理由として考えられるが、タッチスクリーンを介した操作も理由の一つとして考えられる。

タッチスクリーンは直接画面に触れるため、別の装置を介すといった煩雑な手続きが必要なく [1]、直感的に操作できると考えられる。ゲームパッドとマウスとタッチスクリーンを同一のビデオゲームで比較すると、タッチスクリーンが最も操作しやすいという結果が報告されている [2]。これらはタッチスクリーンが直感的で操作しやすく、スマートフォンで提供されるビデオゲームの内容に適していたため、こうした結果が得られたと考えられる。

¹ 東京大学 学際情報学府
The University of Tokyo, 7-3-1 Bunkyo, Tokyo 113-8656, Japan

^{a)} hakuta-mitsuru521@g.ecc.u-tokyo.ac.jp

ビデオゲームの操作デバイスに関してはタッチスクリーン以外にも研究されており、PlayStation Move[3]といった棒状の操作デバイスは従来のコントローラーや Kinect[4]を利用して身体だけで操作する場合と比べて、直感的な操作に適しているということが報告されている [5]。これは煩雑なボタン操作を要求されるコントローラーや身体全体を駆使して操作する Kinect よりも、操作が単純明快でビデオゲームの内容も複雑な動作がなかったためだと考えられる。実際に、入力方式を選ぶ際には操作デバイスと操作するシステムとの関係を考慮する必要があると報告している [6]。

また卓球のビデオゲームでラケットの形状をした操作デバイスは Kinect を利用して手だけで操作する場合と PlayStation Move といった棒状の操作デバイスを利用する場合と比べて一層現実近く、自然であると感じるため、操作することが容易であると報告されている [6]。つまりビデオゲームの内容が現実に存在するのであれば、実際に使用する物体に近い操作デバイスにすることで自然に扱いやすくなると考えられる。

本稿では、我々が開発した既存のビデオゲームを元に新たなデバイスを追加することで、人間とシステムとのインタラクションを円滑に行える新たなゲームの形態を提案する。

2. 作品設計

2.1 セレブラボウリング

まず我々が開発した既存のビデオゲーム「セレブラボウリング」について説明する。

2.1.1 開発の経緯

「セレブラボウリング」は我々が元々所属していた研究室が「世界一受けたい授業」というテレビ番組に出演する際に、開発されたビデオゲームである。本ゲームはゲームエンジン Unity3D[7] で開発されている。番組で紹介された本ゲームは好評であり、合計 3 回同番組に出演し、紹介された (2015 年 7 月 25 日放送回, 2015 年 9 月 26 日放送回, 2015 年 11 月 28 日放送回)。その度に新たな問題を作成したり、画面遷移の表現を変更するといったブラッシュアップを行った。

2.1.2 ゲーム内容

本ゲームでは、図 1 のように 10 本のボウリングのピンが並んでいる画面が表示される。ボウリングのピンには各々異なったひらがな 3 文字の単語が貼りつけられている。ひらがな 3 文字のうち特定の 1 文字が虫食い状態になっている。この虫食い状態の箇所全てに同一で特定のひらがな 1 文字を当てはめると、10 ピン全てに貼り付けられている単語が成立する。このひらがな 1 文字を考える、という虫食いクイズのゲームになっている。成立した単語につきボウリングの演出が施される仕様である。

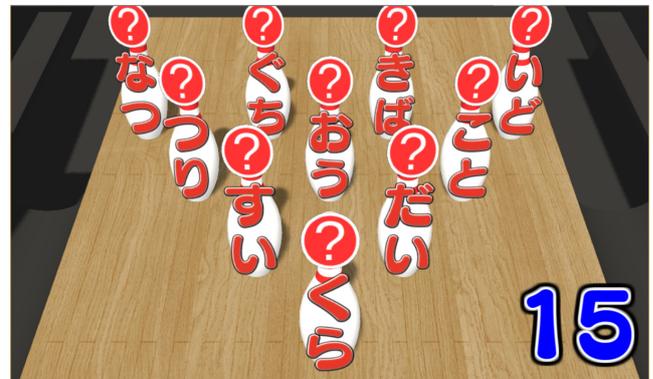


図 1 ゲーム画面
Fig. 1 Game screen

図 1 であれば「か」を当てはめると「花王」「課題」「加水」という言葉が成立し、ボウリングの演出としては 3 ピン倒れる演出がなされる。正解となるひらがな 1 文字は表示された画面に対し唯一なため、10 ピン全てに貼り付けられている単語が成立し、ストライクの演出がなされるのは特定の 1 文字だけである。

2.1.3 ゲームの流れ

ゲームの流れとしてはまずタイトル画面が表示され、次に問題画面が表示される (図 1)。この画面は 15 秒間表示され、その後解答入力待ちの画面となる。解答入力待ちの画面では虫食い状態の箇所どのひらがなが正解となるのか考え、その 1 文字のひらがなを入力する (図 2)。解答を確定すると並んでいるピンに向かって入力されたひらがな 1 文字を乗せたボールが転がっていくボウリングの演出がなされる (図 3)。入力されたひらがな 1 文字によって成立する単語分だけピンが倒れる。その後結果表示画面となり、倒れたピンの本数、成立した単語が表示される (図 4)。

2.2 ボール型デバイス

本研究では 2.1 節で紹介した「セレブラボウリング」を元に、ボール型デバイスを新たに開発した (図 5)。本デバイスは直径 19.4cm の半球状の発泡スチロールを 2 つ組み合わせ構成されている。ボールの外見はゲーム画面に表示されるボールと同じく、緑色で着色してある。ボウリングのボールらしく把持できるよう 3 箇所穴が空けてある。ボール型のデバイスを使用した本ゲームを「コトダマボウリング」と命名した。ボール型である理由は、「セレブラボウリング」がボウリングをモチーフにしたビデオゲームであり、現実で扱うボールと同じ形状であることで自然に扱



図 2 解答入力
Fig. 2 Answer input



図 4 結果画面
Fig. 4 Result screen



図 3 ボウリング演出
Fig. 3 Bowling expression

いやすく、直感的に操作できると考えたためである。

そのため本デバイスには加速度センサーが搭載されている。

2.2.1 加速度センサー

本デバイスにおいて加速度センサーは解答と思われるひらがなを確定(図 2)させた後、ボウリングの演出をする際に用いられる。内部の回路は図 6 で示される Arduino Nano[8] と加速度センサー、Bluetooth モジュールを用いて構成されている。加速度センサーによる 3 軸の加速度を検知した信号を Arduino Nano へと送信し、一定以上の加速度を検知した場合のみ Bluetooth モジュールを介して

ゲームシステムに送信される。これにより図 2 の画面でボール型デバイスを画面の方向に振ることで、手に持っているボール型デバイスがそのまま画面内部に入って転がっているかのような自然なインタラクションが実現できると考えた。

2.2.2 マイク

本デバイスには Bluetooth マイクも搭載されている(図 2)。マイク入力の妨げにならないようにボール型デバイスのマイクの入力に当たる部分には細かい穴が 7 つほど空けてある。マイクには ELECOM LBT-HS10 を使用した。本マイクは解答と考えられるひらがな 1 文字の入力を行う。

これは音声認識を用いた入力となる。

2.2.3 音声認識

昨今、スマートフォンにおける Siri[9] やしゃべってコンシェル [10] といったユーザーアシスタント、Google 音声検索や Google ドキュメントにおける音声入力といった音声認識を利用したシステムは生活の至る所で散見される。Siri[9] やしゃべってコンシェル [10] では、スマートフォンに搭載されたメールやカメラといったアプリケーションの起動や電話発信、音楽再生といった処理を本体を触ることなく実行できる。また Google 音声検索や Google ドキュメントにおける音声入力では、文字入力のためにキーボードを用いて入力したり、スマートフォンにおけるタッチスクリーンによる入力を行う必要がなく、音声認識された言葉、文章を即座に検索でき、即座にテキストデータに変換できる。これらの機能は両手が塞がっていて本体を触ることができない場合や煩雑な操作が必要となる場面で、画面を触らず他の操作の妨げにならずに円滑に処理できる。



図 5 ボール型デバイス
Fig. 5 Ball type device

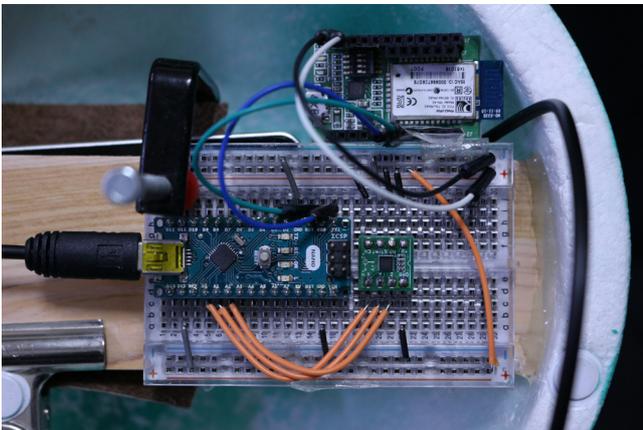


図 6 加速度センサー回路
Fig. 6 Accelerometer circuit



図 7 Bluetooth マイク
Fig. 7 Bluetooth microphone

これらの特徴からボール型のデバイスにボタンを付けて解答を入力するよりも煩雑でなく、円滑に入力できると考え、Bluetooth マイクを搭載した。マイクに入力された音声は Bluetooth 通信でゲームシステムへと送信される。送信された音声は Julius[11] というオープンソースの音声解析ソフトを通じ、解析され、ゲームシステムに入力される。

3. 展示

本作品を用いて3つの会場で展示を行った。そこで得られた感想を記述する。

3.1 東京大学制作展

本作品は第17回東京大学制作展（2015年11月12日～2015年11月16日）において五日間展示を行った。来場者に遊び方を説明し、実際に体験して頂いた。

画面に表示される問題の構成について興味がある体験者や正解が判明するまで繰り返し体験する体験者、テレビ番組で放映されていたことを認知している体験者といったように、ゲームシステム自体の感想が主なものであった。展示初日から周りの展示物から発せられる音や会場で話されている声によって音声認識の精度は悪く、翌日からは解答入力の部分は我々が手動で入力した。

3.2 あそぶ！ゲーム展

本作品は2015年10月3日から2016年2月28日まで埼玉県川口市にあるSKIPシティ彩の国ビジュアルプラザ映像ミュージアムにて開催されていたあそぶ！ゲーム展においても体験して頂く機会を得た。2016年1月10日に開催された特別ワークショップ Vol.3『デジタルゲームを、もっと楽しく！面白く！』というイベント内にて親子40名を対象に本作品を体験して頂いた。本ワークショップでは1台のディスプレイにゲーム画面を投影し、子供たち全員が画面を見えるよう配慮した。

今回は本作品を体験して頂いた後に、アンケートの記入のご協力をお願いした。結果として12名のアンケートを収集できた。アンケート項目ではポジティブな面とネガティブな面を両方記入するようお願いした。本アンケートでは「新しい言葉を知れた」「言葉を考えることが面白かった」といったポジティブな意見が得られたものの、ネガティブな意見としては言葉が難しすぎる点や殆どの回答で音声認識の精度が悪い点について記述されていた。

3.3 第2回バンダイナムコエンターテイメント大運動会

本作品は2016年7月2日に東京ドームで開催された第2回バンダイナムコエンターテイメント大運動会でも展示された。大運動会はバンダイナムコエンターテイメントの社員とその家族が参加するイベントで、玉入れや綱引きといった種目が東京ドームの広大な敷地を使用して行われた。それらの種目とは別に子供用の落書きスペース、折り紙などの工作スペース、また身長を Kinect を用いて測定する装置など来場者が一人でインタラクティブに体験できるスペースも存在した。

我々の作品もそのスペースで展示を行った。本イベントでは1日で223名の来場者の方に体験して頂いた。体験し

た方の多くは子供であり、中には2, 3度体験する来場者もいた。

4. 考察

ボウリングを模した虫食いクイズというゲームシステム自体やボール型のデバイス自体は評判が良いもので、体験者の自然で直感的な操作ができていると考えられる。それはボウリングという内容に則した形状をした操作デバイスであることが理由だと考えられる。

しかし、各展示において音声認識の精度の悪さが目立つ結果となった。この精度の悪さには3点の理由があると考えられる。

1点目は1文字音声認識ということの難しさである。通常の音声認識では、日本語特有の文の流れや各単語におけるひらがな1音の生起確率などから推測する。しかし1文字の音声認識では上記の理論が通用せず、完全に1音だけで判断しなくてはならない。そのため「う」と「る」のといった母音が同一なもの、母音同士の違いなどが認識されない。

2点目は展示会場における雑音である。基本的に音声認識は静かな場所で行うことを前提としており、雑音が多くある場所ではその効果を発揮し得ないと考えられる。各展示会場では他の展示物の音が発生していたり、他の来場者の話し声も観測される。それらが雑音となって認識を妨げていたと考えられる。

3点目は発泡スチロールで形成したボール型デバイス本体の問題である。発泡スチロールという素材は少し触っただけでマイクにとっては大きな雑音を生じさせてしまう。またマイク用に空けていた穴はそれほど音声を通させているわけではなく、寧ろ籠もった音声をマイクに入力させてしまう。更にボール内部では音声が反響しあい、認識を妨げていたと考えられる。

5. まとめ

本研究ではクイズゲームとして成立している「セレブラボウリング」にボール型のデバイスを追加し、新たに「コトダマボウリング」と命名し、3つの展示会場で体験して頂いた。音声認識と加速度センサーを用いることで現実的で扱いやすく自然に体験できる新たなゲームを目指した。結果としてボール形状と加速度センサーは意図通りに機能したが、音声認識の精度が悪く、その点を不満に感じる声が多かった。そのため目的が達成できたとは言い難く、音声認識については改善の必要がある。音声認識ソフトウェアのJulius自体のチューニングや球体形状を形どる素材の変更などで改善することが予想される。また1文字音声認識専用のエンジンを構築することで精度を向上させることも可能だと考えられる。

今回は、展示を行い、そこで得られた感想のみの評価と

なっているため、ボール型デバイスを用いることで意図通りのビデオゲームとして成立しているか実験行って確認する必要がある。今後は音声認識の改善と実験を中心に進めていく。

参考文献

- [1] Motti, L. G., Vigouroux, N., and Gorce, P. Interaction techniques for older adults using touchscreen devices: a literature review. In *Proceedings of the 25th Conference on l'Interaction Homme-Machine IHM '13*, ACM Press (2013), pp. 125-134.
- [2] 國富彦岐, 石川晃, 田所康隆, 白井暁彦. 年齢層とゲーミングデバイスの違いによる面白さの比較調査, エンタテインメントコンピューティングシンポジウム 2013 論文集, pp. 111-116
- [3] PlayStation(R)Move モーションコントローラー — プレイステーション (R) オフィシャルサイト (online), <http://www.jp.playstation.com/ps4/peripheral/cechzcml1jy.html> (2016.12.22).
- [4] Xbox One Kinect センサー (online), <http://www.xbox.com/ja-JP/xbox-one/accessories/kinect-for-xbox-one>
- [5] Birk, M., and Mandryk, R. L. Control your game-self: effects of controller type on enjoyment, motivation, and personality in game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, ACM Press (2013), pp. 685-694.
- [6] Freeman, D., Hilliges, O., Sellen, A., O' Hara, K., Izadi, S., and Wood, K. The role of physical controllers in motion video gaming. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, ACM Press (2012), pp. 701-710.
- [7] Unity - Game Engine(online), <http://unity3d.com/jp/> (2016.12.22).
- [8] Arduino - Home(online), <https://www.arduino.cc/> (2016.12.22).
- [9] iOS - Siri - Apple (日本) (online), <http://www.apple.com/jp/ios/siri/> (2016.12.22).
- [10] シャベってコンシェル — サービス・機能 — NTT ドコモ (online), https://www.nttdocomo.co.jp/service/shabette_concier/ (2016.12.22).
- [11] 大語彙連続音声認識エンジン Julius(online), <http://julius.osdn.jp/> (2016.12.22).