

GhostChord：楽器演奏波形を HID キーボード入力に変換する外部デバイスと音楽ゲームへの応用

粟村 昌弘¹

GhostChord: Converting Real-Instrument Audio into HID Keyboard Input and Its Application to Music Learning Games

1. はじめに

近年、実在の楽器演奏とデジタルゲームを組み合わせた学習支援が注目されている。Fender 社の調査では、ギター初心者が多くが1年以内に離脱するとされ [4]、パンデミック期には約 1,600 万人が新たにギターを始めたと報告されている [5]。また、初心者の約半数は女性であるとの報告もあり [6]、従来の「ロック・ギターヒーロー像」を前提とした市場構造が女性や多様なプレイヤーを周縁化してきたことが指摘されている [7]。

音楽教育の効果は認知発達や社会性の面からも広く示されており [8]、多様な学習者が継続的に楽器演奏に取り組める環境づくりが求められている。一方で、実在楽器対応型の音楽ゲームは専用コントローラのコストや在庫、ゲームタイトル終了後の「周辺機器疲れ」(peripheral fatigue) によって継続利用性が損なわれやすい [1], [2]。

本研究では、実在楽器から得られる音声信号を外部デバイスで解析し、標準的な HID キーボード入力へ変換する汎用インタフェース「GhostChord」を提案する。ゲーム側には通常のキーボードとして認識させることで、特定のゲームエンジンや専用周辺機器に依存しない実装を可能にすることを目的とする。

2. 関連研究・背景

2.1 実在楽器と音楽ゲーム

Guitar Hero[9] や Rocksmith[10] は、実在楽器や専用コントローラを用いる代表的な音楽ゲームである。しかし、周辺機器の製造コストやライセンス料の高騰によりシリーズ終了に至ったことが指摘されている [2]。特に Rock Band 3 以降は、専用コントローラが「一過性のブーム後に使われなくなる周辺機器」としてユーザに認識され、peripheral fatigue を招いたと報告されている [1]。

文化人類学的観点からは、Guitar Hero / Rock Band による「バーチャルな演奏熟達」がいかに体験されるかが分析されている [11]。また、音楽ゲームプレイヤーのリズム能力や音楽性が一般人口と比較してどのように異なるかを検証した研究も存在する [12]。これらの先行研究は、音楽ゲームが学習や演奏体験に与える影響を示す一方で、実在楽器との接続方法については十分に一般化されていない。

2.2 音声解析・ピッチ推定

実在楽器入力をゲームで扱うためには、リアルタイムな音高推定が必要となる。基本周波数推定アルゴリズムとしては、YIN [3] や深層学習ベースの CREPE [13] が代表的である。また、音響信号処理の基礎理論やフィルタ設計については Smith による体系的な整理がある [14]。本研究では、弦楽器の豊かな倍音構造に対して比較的頑健に動作し、かつ実装が単純であることから YIN 系手法を採用する。

¹ 東京国際工科大学 工学部 デジタルエンタテインメント学科 (水上恵太研究室・武本充治研究室)
International Professional University of Technology in Tokyo, Faculty of Technology, Department of Digital Entertainment / Keita Mizukami Lab. & Michiharu Michael Takemoto Lab.

2.3 アクセシビリティとマルチモーダル入力

非言語音声によるゲーム操作に関する研究では、声を用いたインタラクションが運動機能に制約のあるユーザに有効であることが示されている [15]. さらに, Xbox Adaptive Controller のように, 既存ゲーム機へ柔軟に接続できる入力デバイスは, アクセシビリティ向上の観点からも注目されている [16].

本研究で提案する「音声→HID キーボード」変換は, 楽器演奏を主対象としつつも, 同様の枠組みを用いて音声駆動インタフェースを構成できるという点で, 教育とアクセシビリティの双方に跨るポテンシャルをもつ.

3. 提案方式: GhostChord

3.1 問題分析と設計要件

実在楽器入力に関する代表的アプローチを (A) ゲームエンジン内部での音声解析, (B) MIDI 楽器の使用, (C) 専用ギター型コントローラ, (D) Rocksmith 型 DSP パイプライン, (E) 外部音声解析+HID 変換 (本研究) の5種類に整理し, 比較した^{*1}.

(A) は実装を一体化できる利点がある一方, 多声解析の計算負荷やノイズ耐性の問題から教育用途に求められる安定性を満たしにくい. (B) は高精度だが, ギター/ベース学習者を排除することになる. (C) は商用ゲームに近い体験を提供できるが, 製造コストと在庫リスクが大きく, 演奏技能の転移も弱い. (D) は高精度だが, 個人開発規模では再実装が現実的ではない.

これらと比較すると, (E) は「標準的な実在楽器を改造なしで利用できる」「解析処理が軽量」「ゲーム側の実装が単純」という点で, 教育用途における実現可能性・安定性・拡張性のバランスが最も良いと判断できる.

本方式の設計要件を以下にまとめる.

- 標準的なエレキギター/ベースを改造なしで利用できること.
- 音声解析は軽量かつリアルタイムに安定して動作すること.
- ゲーム側は OS 標準のキーボード入力のみを扱えばよいこと.
- 厳密な誤り検出ではなく, 音楽的に意味のあるフィードバックを重視すること.

3.2 システム構成と用語

図1に GhostChord の全体構成を示す. 外部デバイスが楽器音声を取得・解析し, 結果を HID キーボード入力として PC へ送る. ゲーム側は通常のキーボードとして入力を受け取るのみであり, 音声解析の詳細を意識する必要はない.

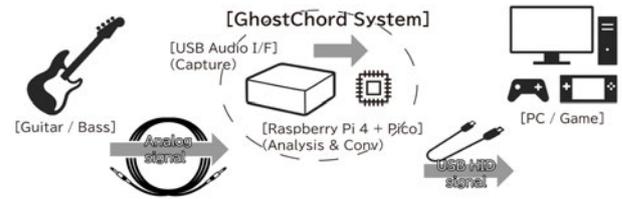


図1: GhostChord のシステム構成

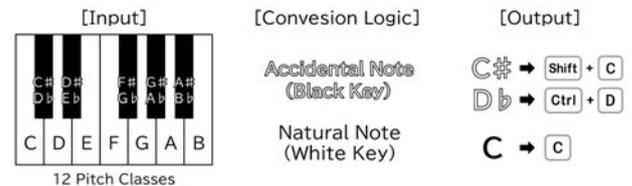


図2: ピッチクラスと HID キーへの割り当て例

本稿では, 以下の用語を用いる.

- 外部デバイス: 楽器から入力された音声をリアルタイムに解析し, HID キーボード入力を生成する独立装置 (本実装では Raspberry Pi 4).
- ピッチクラス: オクターブ差を無視して 12 音 (C, C#, D, ..., B) に分類した音高.
- 構成音加点方式: コード構成音のいずれかが入力された場合に「正解」とみなし, 加点する評価方式.

3.3 音声から HID キーボードへの変換

GhostChord の処理パイプラインは以下のとおりである.

- (1) 標準的なシールドケーブル (1ch 出力) で楽器音声を取得する.
- (2) YIN [3] などのアルゴリズムで基本周波数を推定する.
- (3) 推定音高を最も近いピッチクラスへ量子化する.
- (4) 対応するキーボードイベント (1~2 キー) を生成し, USB HID として送出する.

一般的なエレキギターは 6 弦であり, 同時発音数も最大 6 音である. 一方, HID キーボードには同時押下キー数に関する 6 キー・ロールオーバー (6KRO) 制約があるが, 本方式では多声をそのまま HID へ送るのではなく, 最大振幅成分から 1 音を抽出し, 必要に応じて修飾キーを組み合わせるため, 同時押下数は常に 1~2 キーに収まる. したがって, HID 側の制約と楽器側の物理構造が矛盾することはない.

3.4 ピッチクラスとキー割り当て

図2に, 12 のピッチクラスを少数の物理キーに写像した例を示す. C~B をあらゆる基本キーに加え, シャープ/フラットを修飾キーとして割り当てることで, 10 キー未満で半音階全体を表現できる.

*1 詳細は卒業論文本文を参照されたい.

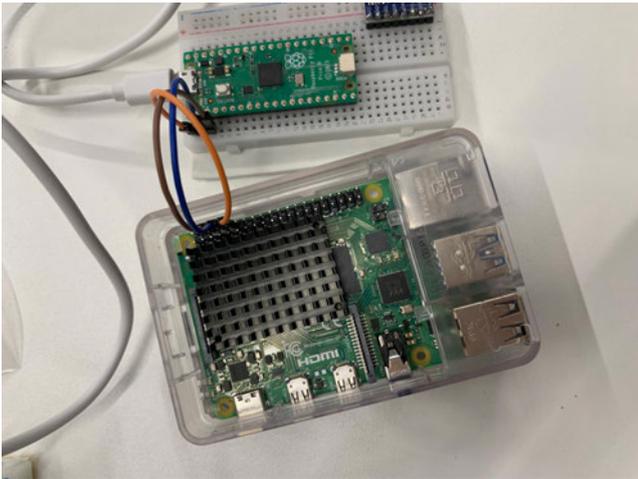


図 3: 外部デバイスの外観



図 4: 接続構成例

3.5 デバイス外観と接続構成

図 3, 図 4 にデバイス外観と接続例を示す。実装では Raspberry Pi 4 に USB オーディオインタフェースと USB ケーブルを接続し, PC 側からは単一のキーボードデバイスとして認識される。

4. 実装とアプリケーション

4.1 HTML/JavaScript によるゲーム実装

初期プロトタイプは Unity を用いて実装したが, 最終的には配布性とアクセシビリティを重視し, HTML/JavaScript による Web 実装へ移行した。ブラウザ側では, GhostChord から送られる入力を通常のキーボードイベントとして扱うのみである。

コード進行は, 拍ごとのコード名と構成音集合として定義する。プレイ中に入力されたピッチクラスが構成音集合のいずれかに含まれていれば加点し, 誤った入力に対して減点は行わない。これにより, 音程の探索過程もゲーム体験の一部として許容し, 初学者に過度な失敗フィードバックを与えない設計とした。



図 5: ジャズ学習向けゲーム画面例



図 6: 音名学習ミニゲーム画面例

4.2 ジャズ学習向けリズムゲーム

図 5 に, ジャズのコード進行に沿って演奏する学習ゲームの画面例を示す。画面上部にコード進行と拍位置を表示し, 下部の譜面に対応するタイミングで任意の構成音を入力することで, スコアと観客のリアクションが変化する。

4.3 音名学習ミニゲーム

図 6 は, 指板上の音名とピッチクラスの対応を学ぶためのミニゲーム画面である。画面上に表示されるターゲット音名に対して, 対応するフレット位置を弾くと魚が釣れるという, 比較的単純なルールとした。GhostChord は同一のデバイス構成のまま, 異なるゲームデザインにも再利用できることが分かる。

5. 評価

本章では, 上級ギタリスト 1 名と成人初心者 1 名を対象とした定性的プレイテスト, および入出力遅延の定量評価について報告する。

5.1 プレイテスト: 上級ギタリスト

1 人目の参加者は大学ジャズ研究会に所属する上級ギタリストである。エレキベースをシールドケーブルで外部デバイスに接続し, ジャズ学習ゲームをプレイしてもらった(図 7)。

観察の結果, 1ch 解析と構成音加点方式の組み合わせにより, プレイ中を通して安定したフィードバックが得られ



図 7: 試遊の様子 (上級ギタリスト)



図 8: 試遊の様子 (初心者)

た。音高推定値が一時的に揺らぐ場面もあったが、多くの場合ターゲットとなるコード構成音の範囲内に留まり、音楽的な妥当性を損なうほどの誤判定は生じなかった。参加者からは「実際のコード進行の練習ツールとして有用である」といったコメントが得られた。

5.2 プレイテスト：成人初心者

2人目の参加者は、それまで楽器演奏経験のない成人初心者である。音名学習ミニゲーム(図6)を中心にプレイしてもらい、理解容易性や学習意欲への影響を観察した(図8)。

参加者は短時間でインタラクションのルールを理解し、数分以内に正しい入力を継続的に生成できるようになった。正しいピッチクラス検出と連動した視覚フィードバックにより、「楽器を学ぶのは楽しい」という肯定的な感想が得られ、セッション後には電子ウクレレの購入を検討し始めた。これらの結果は、GhostChord が初心者の音楽的理解と継続的な練習意欲の双方を支援し得ることを示している。

5.3 入出力遅延の測定 (定量評価)

5.3.1 測定方法

GhostChord の時間的応答性を評価するため、実演奏入力からゲーム画面の判定表示までの end-to-end 遅延を定量的に測定した。実在楽器による演奏を iPhone 17 Pro

表 1: 弦の入力から判定表示まで end-to-end 遅延測定結果

Case	Input Time	Output Time	Frame Diff	Latency [s]	Latency [ms]
1	00:10:25	00:12:26	61	0.254	254
2	01:21:03	01:23:03	60	0.250	250
3	02:11:08	02:13:07	59	0.246	246
4	03:00:03	03:02:12	69	0.288	288
5	03:46:18	03:48:24	66	0.275	275
6	04:26:21	04:28:19	58	0.242	242
7	05:08:16	05:10:15	59	0.246	246

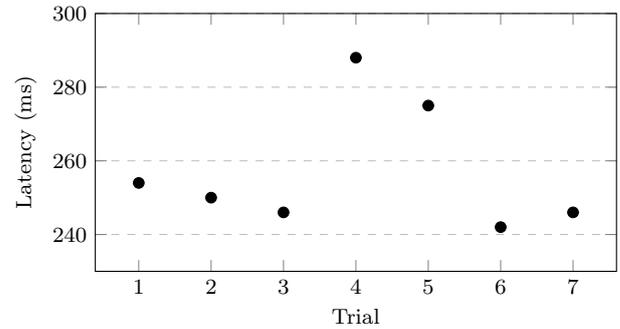


図 9: 各試行の遅延値



図 10: end-to-end 遅延の分布 (箱ひげ図)

の 240 fps スローモーションモードで撮影し、録画データを Adobe Premiere Pro に取り込んでフレーム単位で解析した。

240 fps の映像は 30 fps タイムライン上では 1/8 倍速となるため、1 フレームは実時間の 1/240 秒に相当する。弦を弾いた瞬間と判定表示が出る瞬間のフレーム差を ΔFrames とすると、遅延 Latency は

$$\text{Latency} = \frac{\Delta\text{Frames}}{240}$$

として算出できる。同一条件で 7 回の試行を行った。

5.3.2 測定結果

表 1 にタイムスタンプとフレーム差分、図 9 および図 10 に各試行の遅延値と分布を示す。すべての試行は概ね 240–290 ms の範囲に収まり、平均 257 ms、標準偏差 17.3 ms であった。

5.3.3 考察

測定された 240–290 ms 程度の遅延は、音声バッファリング、基本周波数推定、HID 変換、描画処理などを含む Rocksmith 型の実在楽器入力ゲームにおける 150–250 ms

程度の遅延 [10] と同程度のオーダである。厳密なタイミング精度が重視されるアーケード型リズムゲーム（多くは 50 ms 以下の遅延を目標とする）には不向きである一方、GhostChord はコード構成音に基づくフィードバックを重視しており、250 ms 前後の遅延は学習効果や音楽的意味を大きく損なわないと考えられる。

6. 考察

6.1 アーキテクチャの有効性

プレイテストと遅延評価の結果から、外部音声解析 + HID 変換というアーキテクチャは、実在楽器を用いた学習ゲームにおいて技術的・教育的に両立した解を与えることが示唆された。ゲーム開発者は標準的な入力処理のみを実装すればよく、複雑な DSP パイプラインを各タイトルごとに再実装する必要がない。これは、中小規模の開発チームやインディゲーム開発者にとっても現実的なアプローチである。

6.2 教育的示唆

構成音加点方式により、コード構成音やスケールといった音楽理論上の概念を、そのままゲームのルールとしてマッピングできる。初心者参加者から得られたポジティブな反応は、GhostChord が「楽器学習への入り口」として機能し得ることを示している。音楽ゲームがプレイヤーの音楽性やリズム感に一定の影響を与えることは既存研究でも報告されており [12]、本方式も同様の枠組みの中で位置づけられる。

6.3 アクセシビリティへの応用可能性

本システムは、任意の音声を標準的な HID キーボード入力に変換するだけでよいため、既存ゲーム側には一切の改造を要求しない。音声によるゲーム操作が運動機能に制約のあるユーザの支援に有効であることは既に報告されており [15]、Xbox Adaptive Controller のような汎用入力装置も実用化されている [16]。GhostChord の枠組みを応用すれば、特定の発声や楽器音を「上」「下」「決定」などの入力に割り当て、市販ゲームを音声・楽器駆動インタフェースで操作することも可能となる。

7. まとめ

本稿では、実在のエレキギターおよびベースの音声を外部デバイスで解析し、HID キーボード入力へ変換するインタフェース GhostChord を提案した。音声解析をゲームエンジン外部に切り出し、構成音加点モデルを採用することで、従来の実在楽器対応型ゲームが抱えてきた専用コントローラ依存性や開発コストの問題を緩和できることを示した。

上級ギタリストと完全初心者を対象としたプレイテス

トでは、本方式が音楽的に妥当で教育的にも有用であり、練習意欲の向上に寄与し得ることが確認された。また、end-to-end 遅延の定量評価により、本システムが学習用途として許容可能な応答性を有していることを示した。

著者の知る限り、任意の音声信号を既存ゲームで利用可能な HID キーボード入力へ変換する一般的構造を提案した研究は先行例が少なく、本方式に関する特許出願も進行中である。今後の課題としては、対応可能な演奏テクニックの拡張、複数人セッションモードの設計、教育および福祉分野での大規模なユーザ研究、さらには GhostChord を外部デバイス/ソフトウェア SDK として一般公開し、他作品からも利用可能な形で整備することなどが挙げられる。これらを通じて、実在楽器を用いた音楽ゲームとアクセシビリティ技術の橋渡しを行い、より身体的で表現的なインタラクティブ体験の実現を目指す。

参考文献

- [1] Watton, N.: Looking Back to 2010 and Rock Band 3 - the First Death of Rock Band, *TheXboxHub* (2020). Analyzes the impact of peripheral fatigue (controller fatigue).
- [2] Stuart, K.: Guitar Hero axed: why did the series crash and burn?, *The Guardian* (2011). Discusses licensing fees and manufacturing costs. Accessed: 2025-12-19. <https://www.theguardian.com/technology/2011/feb/09/guitar-hero-axed-why>.
- [3] de Cheveigné, A. and Kawahara, H.: YIN, a fundamental frequency estimator for speech and sounds, *The Journal of the Acoustical Society of America*, Vol. 111, No. 4, pp. 1917–1930 (online), DOI: 10.1121/1.1458024 (2002).
- [4] Mooney, A.: Fender CEO Andy Mooney on the Ultra, MIJ Superstrats and the Future of Guitars, *Reverb News Interview* (2019). Mentions retention/abandonment within the first year. Accessed: 2025-12-19. <https://reverb.com/news/interview-fender-ceo-andy-mooney-on-the-ultra-mij-superstrats-and-the-future-of-guitars>.
- [5] Yahoo Finance: Roughly 16M people learned guitar amid Covid, study finds, *Yahoo Finance Video* (2021). Discusses increased guitar learning during COVID-19 based on Fender-related reporting. Accessed: 2025-12-19. <https://finance.yahoo.com/video/roughly-16m-people-learned-guitar-154618492.html>.
- [6] Wang, A. X.: Fender Finds 50 Percent of New Guitarists Are Women, *Rolling Stone*, (online), available from (<https://www.rollingstone.com/music/music-news/fender-new-guitar-player-survey-737744/>) (2018).
- [7] Bayton, M.: *Frock Rock: Girls and Gigs*, Oxford University Press, Oxford (1998). Analyzes the exclusion of women from guitar culture and retail spaces.
- [8] Hallam, S.: The power of music: Its impact on the intellectual, social and personal development of children and young people, *International Journal of Music Education*, Vol. 28, No. 3, pp. 269–289 (2010).
- [9] Harmonix Music Systems: *Guitar Hero* (2005). Video game, RedOctane Publishing.
- [10] Ubisoft Entertainment: *Rocksmith 2014 Edition - Remastered* (2014). Video game, Ubisoft.
- [11] Miller, K.: *Schizophrenic Performance: Guitar Hero*,

- Rock Band, and Virtual Virtuosity, *Ethnomusicology*, Vol. 53, No. 3, pp. 398–429 (2009).
- [12] Pasinski, A. C., Hannon, E. E. and Snyder, J. S.: How musical are music video game players?, *Psychonomic Bulletin & Review*, Vol. 23, pp. 1553–1559 (2016).
- [13] Kim, J. W., Salamon, J., Li, P. and Bello, J. P.: CREPE: A Convolutional Representation for Pitch Estimation, *Proceedings of the IEEE ICASSP*, pp. 161–165 (2018).
- [14] Smith III, J. O.: *Spectral Audio Signal Processing*, W3K Publishing (2011). Online edition. Accessed: 2025-12-19. <https://ccrma.stanford.edu/~jos/sasp/>.
- [15] Harada, S., Wobbrock, J. O., Malkin, J., Bilmes, J. A. and Landay, J. A.: Voice Games: Investigation into the Use of Non-Speech Voice Input for Making Computer Games Accessible, *Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT 2011)*, Springer, pp. 11–29 (2011).
- [16] Microsoft Corporation: Xbox Adaptive Controller (2018). Hardware device (official page). Accessed: 2025-12-19. <https://www.xbox.com/accessories/controllers/xbox-adaptive-controller>.