

演奏技能評価を目的とした初心者向け楽音解析アプリの試作

堀部 貴紀^{1,a)} 森勢 将雅^{1,b)}

概要: 本研究では、演奏技能評価を目的とした楽音の収録・解析を行うスマートフォンアプリのプロトタイプを検討した。ターゲットは、収録された自身の楽音から演奏技能を評価することが難しい初心者である。収録環境を評価する指標として、リアルタイムに等価騒音レベルを算出し、収録中のクリッピング検知を実装した。音声分析合成システム WORLD を用いて、収録された楽音から基本周波数、パワー、スペクトログラムを抽出し、時系列をグラフで出力した。

1. はじめに

楽器演奏では、楽器ごとにさまざまな技術を求められ、長期的な訓練を必要とするため、演奏者の習熟度に応じてトレーニングすべき演奏技能は異なる。そのため、演奏技能は習熟度の高いプロのトレーナーによる評価が一般的である。これまでも、習熟度を考慮した演奏支援システムは検討されている。ピアノ演奏の初期段階における運指の習熟を高める支援システムでは、打鍵や運指を鍵盤上に直接投影することで直観的に提示した [1]。ギター演奏における習熟度を考慮したタブ譜を自動生成するシステムでは、習熟度に応じて難しいフレーズを楽譜への表示させないことを検討した [2]。プロのトレーナーが担っている演奏技能評価を計算機に置き換えることで、ユーザが単独でトレーニングを進めることができる可能性がある。

ジストニアは、過剰な練習によって演奏者の自発的な運動の制御ができなくなる症状であり [3]、過剰な練習に歯止めをかける必要がある。現在は、ジストニアを計測する手法として、すでにジストニアになった演奏者の楽音から得られる基本周波数の時間的変動をもとにジストニアを定量的に解析する技術は確立されている [4]。計算機による演奏技能評価から、過剰な練習によるジストニアの予測ができるようになると社会的に有用であると考えられる。

演奏技能を評価、支援する手法はこれまでも検討されており、多くはそれぞれの楽器が持つ特性や演奏方法を考慮して得られる特徴量を用いて評価されている [5]。そのうち、特徴量として楽音から得られる基本周波数を用いて、基本周波数の時間変動から奏法を判定する手法が検討されている [6]。

本研究では、演奏技能を評価することを目的とした、楽音の収録・解析を行うスマートフォンアプリのプロトタイプを検討した。一般的に普及しているスマートフォンに搭載されているマイクロフォンを用いて、収録された楽音における特徴量を抽出し、可視化することを目指した。

2. 楽音解析アプリのコンセプト

本プロトタイプは、楽器を安定的に演奏することができ、楽音の技能を評価することが難しい初心者をターゲットとする。利用者は、収録や解析のための特別な機材を必要とせず、スマートフォンのみで完結することを目的とした。そのため、解析には GPU 等を必要とせずスマートフォンでも高速に処理可能な音声分析合成システム WORLD [7] を用いた。本プロトタイプでは、収録された楽音に対して基本周波数とスペクトル包絡を抽出し、基本周波数とパワー、スペクトログラムの時系列を可視化した。

3. 実装

本アプリは、(1) 楽音収録 (2) 楽音解析の 2 段階で構成されている。また、本アプリの特徴として、収録から解析まで一連の流れを利用者が少ない操作で行えるように設計を行った。収録・解析を行うプログラムは、iOS のネイティブ言語 Swift を用いて実装した。

3.1 収録

アプリを起動し「START」ボタンを押下すると、3 秒間のカウントダウン後に楽音の収録が始まる。カウントダウン後に演奏されている 5 秒間の楽音が収録される。楽音収録は、Apple が提供している AVFAudio フレームワーク [8] の AVAudioRecorder クラス [9] を用いて行われる。収録された楽音は、量子化ビット数 16 bit、サンプリング周波数

¹ 明治大学

^{a)} cs222034@meiji.ac.jp

^{b)} mmorise@meiji.ac.jp

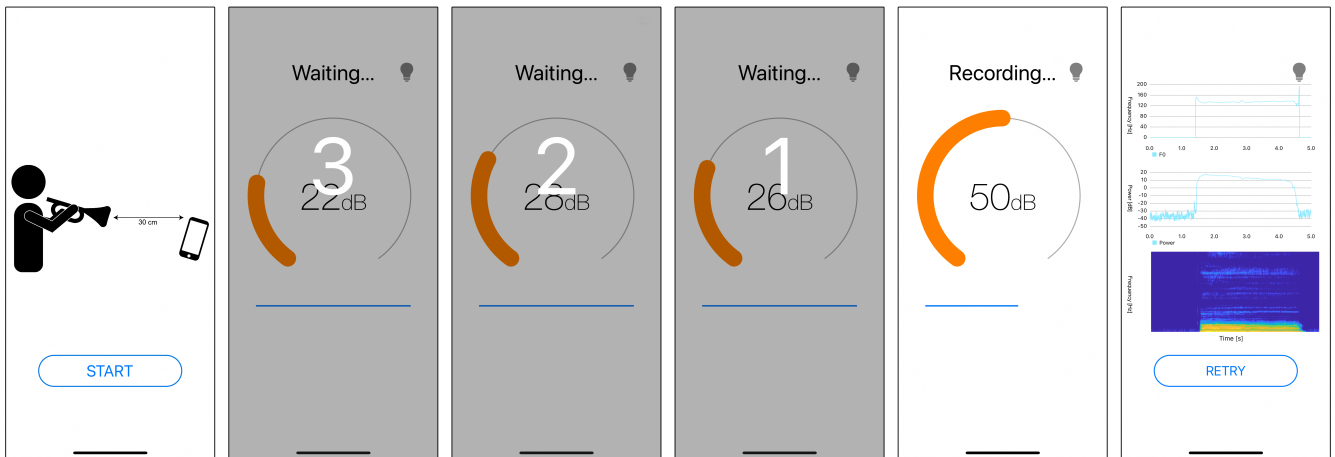


図 1 アプリデザイン

48 kHz, 非圧縮 PCM 形式で保存される。

楽音収録段階では、収録環境と楽音の品質を把握するためリアルタイムに等価騒音レベルの算出とクリッピング検知を実装した。文献 [10] より、等価騒音レベルは 1/3 オクターブバンド分析を利用してベクトル領域で A 特性を与える方式で実装した。得られる値は、キャリブレーションをしていない相対的な等価騒音レベルである。1/3 オクターブバンド分析より得られるそれぞれの帯域における対数パワーに補正值を与えることで、A 特性補正後の対数パワー L_n が得られる。その後、全帯域においてデジタル値から線形値に戻し、総和を求めて再びデジタル値を計算することで、近似的に等価騒音レベル L_{Aeq} が得られる。数式にすると以下の通りである。

$$L_{Aeq} = 10 \log_{10} \left(\sum_{n=1}^N 10^{L_n/10} \right) \quad (1)$$

クリッピング検知では、簡易的な実装として、楽音に対して振幅が 0.95 を超えるときクリッピングしたと判定し、アプリ内のランプを点灯させる。クリッピングしたと判定された収録に関しては、楽音解析画面に遷移するが、楽音解析を行わず、利用者に再収録を促す。

3.2 解析

楽音解析段階では、収録した楽音から抽出した 3 つのグラフを描画する。WORLD における基本周波数推定法である Harvest [11] から得られた基本周波数の時間的変動を描画する。また、WORLD におけるスペクトル包絡推定法である CheapTrick [12] から得られたスペクトル包絡を用いて、収録した楽音におけるデシベル単位のパワーとスペクトログラムを描画する。

3.3 アプリの展望

本プロトタイプでは、スマートフォンで完結する解析のみを行った。今後はサーバとの通信を可能にし、GPU を

利用したアルゴリズムによる解析からユーザーにより深い演奏技能評価をフィードバックすることを検討している。

4. おわりに

本研究では、演奏技能を評価する手法として、スマートフォンで楽音の収録・解析を行うことができるプロトタイプを検討した。収録機能により収録された 5 秒間の楽音から基本周波数、パワー、スペクトログラムの時系列を描画した。今後の展望としては、先行研究 [4] において定義された時間的に変化する基本周波数の揺らぎを抽出し、ジストニアの判定を行う特徴量として、パフォーマンス予測に用いることを検討したい。また、その処理がスマートフォンで高速に処理しない場合、収録した音源をサーバに転送し、サーバ上の GPU による解析を行うシステムの開発を行いたい。

参考文献

- [1] 竹川佳成, 寺田努, 塚本昌彦: 運指認識技術を活用したピアノ演奏学習支援システムの構築, 情報処理学会論文誌, Vol. 52, No. 2, pp. 917-927 (2011).
- [2] Yazawa, K., Itoyama, K. and Okuno, H. G.: Automatic transcription of guitar tablature from audio signals in accordance with player's proficiency, *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3122-3126 (2014).
- [3] Eckart, A.: Focal dystonia: advances in brain imaging and understanding of fine motor control in musicians, *Hand clinics*, Vol. 19, No. 3, pp. 523-538 (2003).
- [4] Lee, A., Furuya, S., Morise, M., Iltis, P. and Altenmüller, E.: Quantification of instability of tone production in embouchure dystonia, *Parkinsonism & Related Disorders*, Vol. 20, No. 11, pp. 1161-1164 (2014).
- [5] 下尾波輝, 矢谷浩司: エレキギター演奏自動評価のための音響的特徴量の調査, 研究報告音楽情報科学 (MUS), Vol. 2017-MUS-117, No. 3, pp. 1-6 (2017).
- [6] 宮垣拓弘, 西野隆典, 成瀬央: 基本周波数の時間変動を用いたエレキギターのピブラト奏法の判定, 研究報告音楽情報科学 (MUS), Vol. 2015-MUS-107, No. 47, pp. 1-3 (2015).

- [7] Morise, M., Yokomori, F. and Ozawa, K.: WORLD: A Vocoder-Based High-Quality Speech Synthesis System for Real-Time Applications, *IEICE Transactions on Information and Systems*, Vol. E99.D, No. 7, pp. 1877–1884 (2016).
- [8] AppleDeveloperDocumentation: AVFAudio, (online), available from <https://developer.apple.com/documentation/avfaudio> (accessed 2022-04-24).
- [9] AppleDeveloperDocumentation: AVAudioRecorder, (online), available from <https://developer.apple.com/documentation/avfaudio/avaudiorecorder> (accessed 2022-04-24).
- [10] 森勢将雅: ひたすら楽しんで音響信号解析, コロナ社 (2021).
- [11] Morise, M.: Harvest: A High-Performance Fundamental Frequency Estimator from Speech Signals, *Proc. Interspeech 2017*, pp. 2321–2325 (2017).
- [12] Morise, M.: CheapTrick, a spectral envelope estimator for high-quality speech synthesis, *Speech Communication*, Vol. 67, pp. 1–7 (2015).