

メロディスロットマシン II: マルチモーダルインタフェースに基づくパート音強調

浜中 雅俊*

概要. 我々はこれまでメロディセグメントを別のバリエーションのメロディセグメントに置き換えることができるメロディスロットマシンというシステムを開発してきた。メロディスロットマシンで複数のパートでメロディを置き換え可能とした場合、音楽初心者はメロディの変化に気が付かない可能性がある。そこで、ユーザのタッチスクリーンへの操作やユーザの頭部方向などのマルチモーダルインタフェースの情報を利用し、ユーザが注目しているパートの音を強調することで、音楽初心者であってもメロディの変化に気づきやすくするメロディスロットマシン II を構築した。

1 はじめに

我々は、メロディセグメントが表示されているダイヤルを回転させることで、メロディセグメントを他のバリエーションに置き換えることができるメロディスロットマシンというシステムを開発した (図 1) [5, 11]。メロディのバリエーションは Generative Theory of Tonal Music (GTTM) に基づくメロディモーフィング法によって生成されているため、メロディセグメントを入れ替えても曲全体の大局的な構造は維持される [7, 4, 3, 6]。演奏者の姿は実写映像がペッパーズ・ゴースト・ディスプレイに映し出されるため、ユーザーはダイヤルを操作することで演奏者を操っているような感覚となる (図 2)。



図 1. メロディスロットマシン



図 2. スロットダイヤル

メロディを聞き取る能力には個人差があり、音楽初心者の場合にはメロディセグメントが置き換わったことに気が付かないことがある。特に複数のパートのメロディが変化する場合には、両方パートのメロディの変化に気が付くには高度な音楽聴取能力が必要である。そこで我々が構築するメロディスロットマシン II では、マルチモーダルインタフェースからの情報を使い、同時に操作可能なパートを1つに限定し、かつ、ユーザが操作しているパートの音を強調することで、メロディの変化を聴き取りやすくすることを目指す。

2 音楽聴取におけるマルチモーダルインタラクション

メロディスロットマシン II では、ボーカルとサックスの2つのパートのメロディセグメントを変化させることが可能である。本研究では音楽聴取においてメロディの変化に気がつくことを目的としているため、楽譜が表示されたダイヤルではなく、横軸が時間、縦軸がメロディのバリエーション (モーフィングレベル) のグリッドパネル状のインタフェースを採用した。ユーザはグリッドパネルを操作しながらそれぞれのメロディを聴くことに集中する。操作可能なパートの限定や、操作しているメロディの強調のために、あらたな新たな操作を追加するとメロディを聴くことに集中することが難しくなりやすい。これに対し、我々が開発したサウンドスコープフォ

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

* 理化学研究所

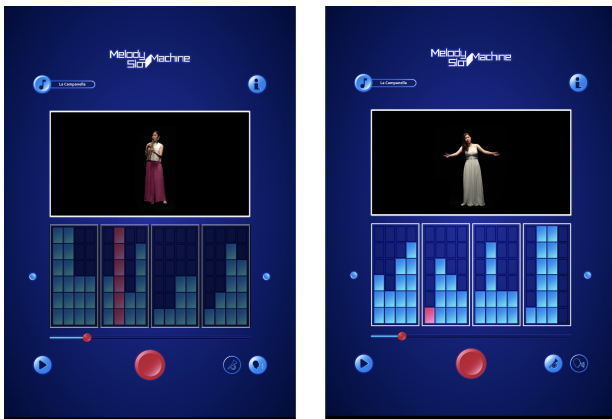


図 3. メロディスロットマシン II

ン・サウンドスコープパッドでは、スマートフォンのフロントカメラでユーザの頭部方向を検出することで、特定のパートの音の強調を可能としていた(図 4) [2, 1, 9, 10]. また、手を耳に近づけて耳を澄ますポーズをするとさらにそのパートの音が強調されるようになっていた. そこで、メロディスロットマシン II では、上記の強調機能を導入する. すなわち、フロントカメラでユーザの頭部方向を検出し、頭部が向いているパートの音を強調する. 耳を澄ますポーズをすると、さらにそのパートを強調する. さらに、ユーザが向いているパートのグリッドパネルはアクティブ、向いていないパートのパネルはインアクティブとなる (図 3).

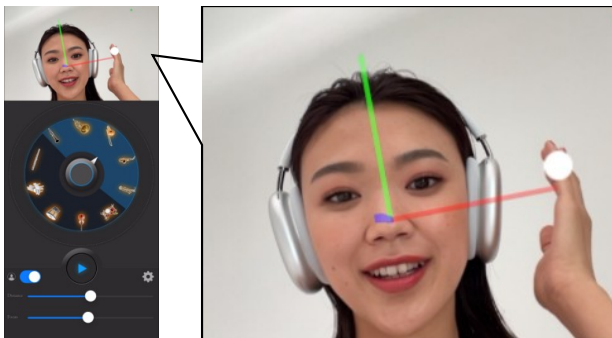


図 4. サウンドスコープフォン

3 実装

ボーカルもサックスもあらかじめモーフィングレベルごとにレコーディングし撮影している. ユーザの操作によってモーフィングのレベルが変化するところでは、別のレベルの映像が結合される. 別のレベルの映像を直接結合すると奏者の姿勢がずれるため不自然となる. そこで我々は、滑らかな動きとなるよう、結合の前と後の 10 フレームの映像を削除し、AI により合成した補間映像を挟み込んだ [8].

メロディスロットマシンは iPad2 台と Mac mini で構成されていて、そのうち一台の iPad が上記の映像の接続処理を行っていた. ボーカルとサックスの 2 つのタイムラインの映像処理を行うには、1 台の iPad のデバイスリソースでは不可能であった. そこで、ボーカルとサックスのすべてのモーフィングレベルの組み合わせの映像 80G バイトのサイズを作成して、1 つのタイムラインで処理を行った.

一方、メロディスロットマシン II では、ボーカルを表示する iPad からサックスを表示する iPad へ同期信号を送ることで、2 つのタイムラインの同期を実現した. 性能の低い iPad のモデルでも実行可能とするため、機能を 2 つの iPad にアサインした. すなわち、ボーカルを表示する iPad でフロントカメラによる頭部方向の検出を行い、サックスを表示する iPad でサウンドの出力を行った.

ユーザが iPad を操作中には、操作している iPad のグリッドタイルがアクティブ、操作していない iPad のグリッドタイルがインアクティブになる. ユーザが操作していない時には、iPad が取得したユーザの頭部方向が正面から 30 度未満の場合にボーカルをアクティブ、30 度以上の場合にはサックスをアクティブとした. 角度の初期値は 30 度であるがスライダーにより調整可能である. 我々はサウンドスコープフォン・サウンドスコープパッドと同じ手法を実装しパートの強調を実現した [2, 1]. パートの強調は、機能のオンオフを切り替えることができる.

4 実験結果

5 名の音楽初心者がメロディスロットマシン II のパートの強調機能の ON と OFF を繰り返しながら 30 分程度試用した後に意見を聞いたところ全員がパートを強調したほうがメロディが聴き取り安いと答えた. 音はポータブルスピーカから出力した. グリッドパネルのアクティブとインアクティブの切り替え機能は必要ないという意見が多かったが、操作をしていない人が操作をしている人を見た場合に、何が行われているのか理解しやすいという意見もあった.

5 おわりに

2 つのパートのメロディを置き換え可能なアプリケーションメロディスロットマシンにおいて、音楽初心者が複数のメロディの変化に気が付きにくい問題を解決するためメロディスロットマシン II を構築した. 実験では、ユーザの頭部が向いている側のパートを強調することによって、5 人の音楽初心者の全員がメロディを聴き取りやすくなった. メロディセグメントの置き換えを 3 つ以上のパートで可能な曲を制作予定である. 今後、メロディスロットマシン II を Apple App Store で公開する予定である.

謝辞

本研究の一部は JSPS 科研費 21H03572 の助成を受けたものです。

参考文献

- [1] M. Hamanaka. Sound Scope Pad: Controlling a VR Concert with Natural Movement. In *Proceedings of the 2022 International Conference on Multimodal Interaction, ICMI '22*, p. 730–732, New York, NY, USA, 2022. Association for Computing Machinery.
- [2] M. Hamanaka. Sound Scope Phone: Focusing Parts by Natural Movement. In *ACM SIGGRAPH 2022 Appy Hour, SIGGRAPH '22*, New York, NY, USA, 2022. Association for Computing Machinery.
- [3] M. Hamanaka, K. Hirata, and S. Tojo. Implementing "A Generative Theory of Tonal Music". *Journal of New Music Research*, 35(4):249–277, 2006.
- [4] M. Hamanaka, K. Hirata, and S. Tojo. Implementation of melodic morphing based on generative theory of tonal music. *Journal of New Music Research*, 51(1):86–102, 2023.
- [5] M. Hamanaka, T. Nakatsuka, and S. Morishima. Melody Slot Machine. In *ACM SIGGRAPH 2019 Emerging Technologies, SIGGRAPH '19*, New York, NY, USA, 2019. Association for Computing Machinery.
- [6] K. Hirata, S. Tojo, and M. Hamanaka. *Music, Mathematics and Language*. Springer Nature Singapore, 2022.
- [7] F. Lerdahl and R. Jackendoff. *A generative theory of tonal music*. The MIT Press, Cambridge, MA, 1983.
- [8] T. Nakatsuka, Y. Tsuchiya, M. Hamanaka, and S. Morishima. Audio-Oriented Video Interpolation Using Key Pose. *International Journal of Pattern Recognition and Artificial Intelligence*, 35(16), 2021.
- [9] 浜中 雅俊. サウンドスコープフォン. 情報処理学会 音楽情報科学研究会 研究報告 2021-MUS-131, No. 61, 3 Pages, jun 2021.
- [10] 浜中 雅俊. サウンドスコープパッド. 情報処理学会 音楽情報科学研究会 研究報告 2022-SLP-142, No. 11, 5 pages, 2022.
- [11] 浜中 雅俊. メロディスロットマシン HD. 情報処理学会 音楽情報科学研究会 研究報告 2023-MUS-137, No. 13, 7 pages, 2023.