

インタラクション可能な仮想演奏者

浜中雅俊¹

概要: 仮想演奏者の演奏を操作する体験ができる、メロディスロットマシンについて述べる。

Interactive Virtual Performer

MASATOSHI HAMANAKA^{†1}

Abstract: We developed an virtual performer called the "Melody Slot Machine," which provides an experience of manipulating a music performance.

1. はじめに

本稿では、インタラクティブな音楽システム、メロディスロットマシンについて紹介する[1,2]. メロディスロットマシンは、複数のバリエーションのメロディがセグメントに分割されていて、そのセグメントがダイヤル型のユーザインタフェースに表示されている。そして、そのダイヤルを回転させることでホログラフィックディスプレイに表示された仮想演奏者が未来に演奏するメロディを選択することができる(図1)。メロディのバリエーションは、LerdahlとJackendoffにより提案された音楽構造解析 GTTM (Generative Theory of Tonal Music)に基づいて構成されているため、メロディセグメントを切り替えても、メロディの全体的な構造が保持され曲が破綻することはない[3].

我々は、これまで15年にわたり GTTM に基づく音楽構造分析器の構築を行ってきた[4-6]. GTTM は、理論自体のあいまいさに加え、音楽の解釈があいまいさが含まれるという問題があり、精度の高い分析器を作ることは困難であった。近年、深層学習の導入により、分析精度が飛躍的に向上し、分析結果を用いた実用的なアプリケーションの構築が期待されている[7-9]. GTTM の分析の結果獲得されるタイムスパン木は、メロディ中の構造的に重要な音符と装飾的な音符との関係を示している。したがって、構造的に重要な音符を残しながら装飾的な音符を段階的に減少させることで、簡約したメロディを抽出することが可能である。2つの入力メロディから中間メロディを生成するメロディモーフィング手法は、この簡約の仕組みを用いたものである[10,11].

2. メロディモーフィング手法

メロディスロットマシンは、初心者がメロディモーフィング手法で生成されたメロディを切り替える操作を体験できるシステムです。

図5は、タイムスパン木を用いたメロディの簡約の例である。図のメロディAの上にある木構造は、メロディAをタイムスパン簡約した結果得られたタイムスパン木である。タイムスパン木のレベルBより下にある枝の音符を簡約(省略)するとメロディBのようになる。さらに、レベルCより下にある枝の音符を簡約するとメロディCのようになる。このとき、メロディBはメロディAとCの間のメロディであることから、メロディの簡約もメロディのモーフィングの一種と考えることができる。

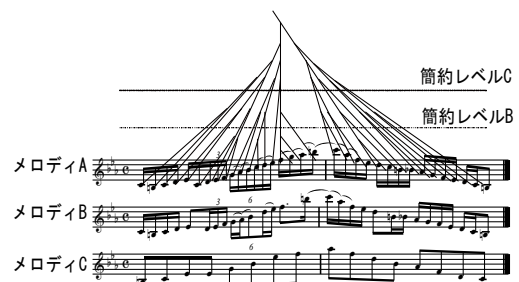


図1 メロディの簡約

メロディモーフィングを実現するため、文献[10-13]で定義されている演算である、包摂関係 \sqsubseteq , meet(最大下界) \sqcap とjoin(最小上界) \sqcup を用いる。包摂関係 \sqsubseteq は、F1を下位の構造、F2を上位の構造(下位の構造を含んでそれ以上の構造を持つ)としたとき、 $F1 \sqsubseteq F2$ と表記し、F2はF1を包摂すると言う。たとえば、図1のメロディA, B, Cのタイムスパン木 T_A, T_B, T_C の包摂関係は、以下のように表せる。

$$T_C \sqsubseteq T_B \sqsubseteq T_A$$

meet(最大下界)は、 T_A, T_B の共通部分のタイムスパン木 $T_A \sqcap T_B$ である。join(最小上界)は、メロディA, Bのタイムスパン木 T_A, T_B が矛盾を起こさない限り統合したタイムスパン木 $T_A \sqcup T_B$ である。

モーフィングされたメロディは、以下の3ステップで獲得できる。まず、2つのメロディA, Bのタイムスパン木 $T_A,$

¹ 理化学研究所
RIKEN

T_B を求め、その共通部分(最大下界) $T_A \cap T_B$ を求める。これにより、タイムスパン木 T_A, T_B は、それぞれ共通部分と非共通部分に分けることができる。

次に、メロディの部分簡約手法を用いて、 $T_A \cap T_B$ と T_A の中間にあるメロディ C を求める。同様に、 $T_A \cap T_B$ と T_B の中間にあるメロディ D を求める。

最後に、メロディ C と D を統合(最小上界 $T_C \sqcup T_D$)し、モーフィングメロディ E を獲得する。

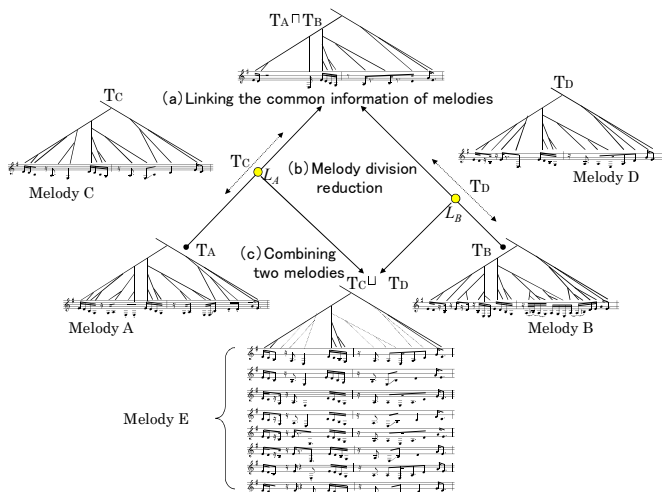


図2 メロディモーフィング手法

3. メロディスロットマシン

誰でも簡単にメロディを操作することを可能とするため、メロディのセグメントの一部を切り替えられるダイヤル型のユーザインタフェースとした(図3)。楽譜を挟んだアクリル板の一部に長方形の穴が開いており、その穴からタブレット上のダイヤル型インターフェースが操作できる。

楽譜の右側にある赤いレバーを下げると、すべてのダイヤルが回転し、ダイヤルのメロディセグメントの1つがランダムに選択されます。演技者の様子を視覚的にも聴覚的にも確認できるようにホログラフィックディスプレイ(ペッパーゴーストホログラム)を使用して出演者を表示し、臨場感が高まるようにした(図4)。

2つのメロディセグメント間でビデオを補間するために、セグメントとセグメントの組み合わせに応じて補間ビデオを生成した[15]。これにより、演奏者の映像の不連続性が解消された。なお、より臨場感を高めたいと考えたため、演奏音は全て収録し録音には、残響が非常に少ないスタジオを使用した。残響があった場合には、セグメントに分割する際に、前のセグメントの残響のみが次のセグメントの先頭に入ってしまう。

4. おわりに

インタラクティブな仮想演奏者を実現するメロディスロットマシンを紹介した。概要映像は以下に公開している。

<https://www.youtube.com/watch?v=DE8j6DtAU5g>



図3 スロットダイヤルとレバー



図4 ペッパーゴーストディスプレイ

参考文献

- [1] Hamanaka, M.: "Melody Slot Machine: A Controllable Holographic Virtual Performer", ACM International Conference on Multimedia (MM'19), 2019, pp. 2468-2477.
- [2] Hamanaka, M., Nakatsuka, T., and Morishima, S.: "Melody Slot Machine", ACM Siggraph, 2019, Emerging Technologies ET-245.
- [3] Lerda, F., and Jackendoff, R. S.: A Generative Theory of Tonal Music, MIT Press, Cambridge, Massachusetts, USA, 1983.
- [4] Hamanaka, M., Hirata, K., and Tojo, S.: Implementing "A generative theory of tonal music," Journal of New Music Research, 2006, 35 (4), 249-277.
- [5] Hamanaka, M., Hirata, K., and Tojo, S.: "ATTA: Automatic Time-span Tree Analyzer Based on Extended GTTM", ISMIR, 2005, pp.358-365.
- [6] Hamanaka, M., Hirata, K., and Tojo, S.: "FATTA: Full Automatic Time-span Tree Analyzer", ICMC, 2007, Vol.1, pp.153-156.
- [7] Hamanaka, M., Hirata, K., and Tojo, S.: "deepGTTM-I: Local Boundaries Analyzer based on Deep Learning Technique", CMMR, 2016, pp.8-20.
- [8] Hamanaka, M., Hirata, K., and Tojo, S.: "deepGTTM-II: Automatic Generation of Metrical Structure based on Deep Learning Technique", SMC, 2016, pp.221-249.
- [9] Hamanaka, M., Hirata, K., and Tojo, S.: "deepGTTM-III: Multi-task Learning with Grouping and Metrical Structures", Lecture Notes in Computer Science, 2018, Vol. 11265, pp. 238-251.
- [10] Hamanaka, M., Hirata, K., and Tojo, S.: "Melody Morphing Method Based on GTTM", ICMC, 2008, pp. 155-158.
- [11] Hamanaka, M., Hirata, K., and Tojo, S.: "Melody extrapolation in GTTM approach", ICMC, 2009, pp. 89-92.
- [12] Nakatsuka, T., Hamanaka, M., and Morishima, S.: "Audio-guided Video Interpolation via Human Pose Features", VISAPP, 2020, Vol.5, pp.27-35, DOI: 10.5220/0008876600270035.