

シリンダーオルゴールを対象とした FCN による自動編曲

松本 優太 酒向 慎司
名古屋工業大学

1 はじめに

オルゴールはその美しい音色によって多くの人に親しまれており、その中で最も一般的なものがシリンダーオルゴールである。シリンダーオルゴールの構造は楽器としては比較的単純なものであり、演奏される楽曲には構造上の制約が課せられる。また、オリジナル曲でない場合は元となる曲の雰囲気を残した編曲を行う必要がある。このような作業には専門知識と経験が必要であり、容易なものではない。

本研究では編曲の手助けとなるよう、シリンダーオルゴールを対象とした自動編曲手法を提案する。具体的には、楽曲間の変換に FCN (Fully Convolutional Network) [1] を適用し、構造的制約を満たした楽曲生成手法を検討する。

2 シリンダーオルゴール

2.1 シリンダーオルゴールの構造

シリンダーオルゴールは、シリンダーとそこに埋め込まれたピン、音が鳴る振動板などによって構成されている。振動板は複数の弁に分かれており、それぞれの弁が一つの音高を担っている。シリンダーが回転し、ピンが対応する弁を弾くことで自動演奏を行う。このように複数の部品を組み合わせて音を鳴らしているが、演奏される楽曲には構造上の制約が幾つか課せられる。本研究ではこれを構造的制約と呼ぶ。

2.2 構造的制約

本研究では、構造的制約のうち重要度の高い弁数条件と音域条件を考慮した編曲を行う。条件を設定するにあたり、田代らのオルゴールの編曲についての解説 [2] や、オルゴール編曲支援ソフト [3] を参考にした。

振動板には一定数の弁があり、その数を弁数と呼ぶ。複数の弁にはそれぞれ特定の音高が設定され、編曲時にその弁数に音を収める必要がある。これを弁数条件と呼ぶ。機種によって弁数は異なるが、本研究では 18 弁のシリンダーオルゴールを対象とする。ただし、これ以降で説明する手法は異なる弁数でも応用可能であると考えられる。

楽曲の弁数を計算する際には、異なる音高の数を 1 弁として数えるが、さらに最少針間隔を考慮する必要がある。弁の振動が収まるまで一定の時間が必要であるため、同一弁に対して連続で鳴らすことのできる間

隔が制限されており、これを最少針間隔と呼ぶ。最少針間隔より短い時間で音を連続で鳴らす場合は、同じ音高の弁をもう一弁追加することで対応する。

また、音域条件は音の高さの範囲を定めるものであり、弁数によって異なる。

3 FCN による自動編曲

シリンダーオルゴールを対象とする自動編曲を目的とした研究は著者が調査した範囲では行われていないが、編曲の一形態であるピアノリダクションの研究が参考になると考えた。

ピアノリダクションとは、ピアノ以外の複数の楽器のために書かれた楽曲の音を適宜削減することでピアノで演奏できるように編曲することである。本研究の目的であるシリンダーオルゴールを対象とした編曲でも、一つの楽器で演奏するために音数を削減するように編曲されており、共通する部分が大きく応用できると考えた。その中でも Hoshi らは、FCN を用いてオーケストラ曲とピアノ曲の関係を学習させることで、ピアノリダクションを行った [4]。この研究と同様に、原曲とオルゴール楽曲の関係も学習できると考えたため、本研究ではこの FCN を利用して自動編曲を行う。

4 提案手法

4.1 編曲モデル

提案する編曲手法を図 1 に示す。Hoshi らの FCN モデル [4] に構造的制約を考慮した機構を加えることで自動編曲を行う。楽曲は音高と経過時間の行列、いわゆるピアノロールで表現する。

弁数条件については、出力された行列成分のうち出力確率の高い方から選択するようなルールを設定することで、18 弁で演奏可能な楽曲を出力する。また、音域条件を満たすために損失の計算を使用可能な音域である 37 音高に限定した。さらに自然さの向上を目的として、原曲のメロディとそれ以外を分けた。入力時には、メロディとメロディ以外の MIDI をそれぞれ 2 次元のピアノロールの行列に変換したものを重ねて、3 次元の配列として扱った。

4.2 データセット

FCN を学習するためには、シリンダーオルゴールの編曲前後のペアデータセットが必要である。しかし、そのようなデータセットは存在しないためデータを収集した。ジャンルはメロディが分かりやすいポップスに限定し、オルゴール楽曲とそれに対応する原曲の MIDI

データのペアを 100 組収集した。オルゴール楽曲については演奏音源を採譜することで MIDI データに変換しており、1 曲の長さは冒頭 14 秒に統一した。また、原曲の BPM、キーはオルゴール楽曲に合わせて変更した。

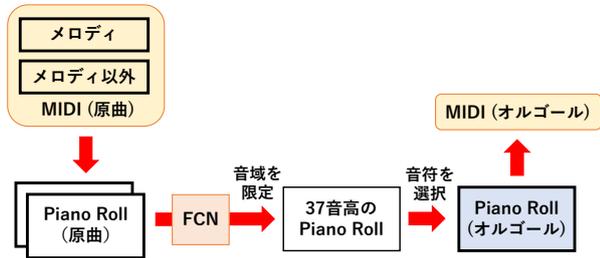


図 1: 提案する編曲手法

5 実験

まずは、編曲モデルの学習を行う。学習させる曲の長さを検討した結果、曲全体では自然に生成できなかったため 1 曲を 4 分割した。また、100 曲 × 4 = 400 個のデータを訓練:テスト = 9:1 として使用した。MIDI データは、128 音階、最小時間単位を $\frac{1}{32}$ 秒として、 $\{0, 1\}^{128 \times 112}$ のピアノロールの行列に変換する。ここで 0, 1 はそれぞれ音が鳴らない、鳴ることを表す。FCN モデルの構造は、入出力データの次元数に関する部分以外は Hoshi ら [4] と同じである。学習時には音域条件に該当する 37 音高の部分のみ平均二乗誤差を計算し、 $\{0, 1\}^{37 \times 112}$ で出力する。その後、4 分割した楽曲を結合した $\{0, 1\}^{37 \times 448}$ の行列に対して選択ルールを適応した。

次に、生成される楽曲の評価を行う。原曲のメロディをそれ以外の音と分けることで自然さが向上するかを確認するため、メロディを分けられない場合と出力結果を比較する。5.1 節、5.2 節ではテストデータ 10 曲の出力結果を用いた。

5.1 客観評価

評価方法

生成楽曲において原曲のメロディが自然に残っているかを確認するために、正解楽曲と出力楽曲の残存メロディ一致率を定義して評価に用いた。ここでは、ピアノロールの行列成分が、両方の楽曲で 1 (音が鳴る) だった場合を一致とみなす。

$$\text{残存メロディ一致率} := \frac{(M_{\text{ori}} \wedge M_{\text{cor}} \wedge M_{\text{out}}) \text{ の音数}}{(M_{\text{ori}} \wedge M_{\text{cor}}) \text{ の音数}} [\%]$$

- M_{ori} : 原曲のメロディ
- M_{cor} : 正解楽曲のメロディ
- M_{out} : 出力楽曲のメロディ

結果と考察

実験の結果、メロディを分けられない場合と分ける場合の一致率はそれぞれ 34.7%, 68.0% となった。これは、メロディを分けることでどの部分を残せばよいかを学習しやすくなったためだと考えられる。

5.2 主観評価

評価方法

自然な楽曲を生成できているかを確認するために主観評価を行う。被験者は 23 名であり、音高の変化、リズム、和音の 3 つの項目について自然さを 5 段階で評価する。流れとして、入力楽曲をオルゴール編曲した正解楽曲を全ての項目の点数が 5 である自然な楽曲として聴いた後、メロディを分ける場合と分けられない場合の 2 つの生成楽曲をどちらか分からない状態で聴いて評価を行う。これを 10 曲分繰り返した。

結果と考察

各曲の平均点の分布を表す箱ひげ図を図 2 に示す。被験者による評価のばらつきはあまり見られなかった。t 検定を有意水準 0.05 で行った結果、全ての評価項目に対してメロディを分ける場合と分けられない場合に有意な差があり、いずれも分けた方が高い点数となった。

項目 1、項目 2 では、メロディが自然なため点数が高くなったと考えられる。項目 3 では点数のばらつきが大きいが、これはメロディ部分で余分な音が生成されなくなった一方で、一部の楽曲で伴奏として同時に鳴らされる不自然な音が増えたためだと思われる。

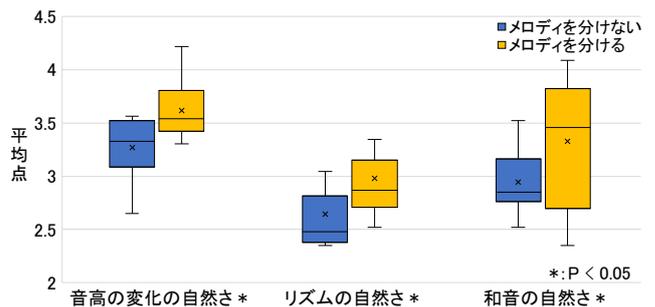


図 2: 主観評価の実験結果

6 むすび

本研究ではシリンダーオルゴールを対象とした自動編曲を行った。構造的制約を考慮した FCN モデルを構築することで、弁数条件と音域条件を満たした楽曲を生成した。また、学習時に原曲のメロディを分けることで自然さが向上することが確認できた。今後の展望として、他の弁数への対応や構造的制約の追加、自然さの向上が挙げられる。

参考文献

- [1] Long, J. et al.: “Fully convolutional networks for semantic segmentation”, Proceedings of the IEEE conference on computer vision and pattern recognition, pp.3431-3440, 2015.
- [2] 田代和夫.: “オルゴールの編曲”, 日本音響学会誌, 45(2), pp.144-149, 1989.
- [3] 原田敬.: “音のキャンバス”, <http://www.01.246.ne.jp/~ttha/>, 参照 2023-01-03
- [4] Hoshi, Y. et al.: “Versatile Automatic Piano Reduction Generation System by Deep Learning”, ICARC, pp.66-71, 2022.