

ピアノ演奏における打鍵の強弱を考慮した MIDI の生成

大谷 紀子 研究室

1572089 向 瑞稀

1. 背景と目的

現在、個人の感性を反映した楽曲の自動生成に関する研究[1]が進められている。入力された既存楽曲に共通する特徴を感性モデルとして獲得し、進化計算アルゴリズムにより個人の感性を反映した楽曲を生成して、MIDI ファイル形式で出力する。MIDI ファイルとは、音量や音高、発音タイミングなど演奏に関する情報が記録されたファイルで、コンピュータ等で再生が可能である。自動作曲システムにより出力される MIDI ファイルは、音高と発音タイミングに関して楽譜を忠実に再現しており、どのように演奏されるかについては考慮されていない。一般的に人間が楽曲を演奏する際には、楽曲を盛り上げるために、発音タイミングをずらしたり抑揚を付けたりする。楽譜を忠実に再現した機械的な演奏では、生成された楽曲に個人の感性が反映されていても、楽曲が盛り上がりには欠け、印象が悪くなると考えられる。本研究では、ピアノ演奏における演奏表情を考慮した MIDI ファイルの生成を目的とし、打鍵の強弱を反映した MIDI ファイルの生成手法を提案する。

2. 打鍵の強弱ルール抽出と反映

MIDI ファイルには、ベロシティとよばれる制御情報が記録されており、楽器演奏における打鍵の強弱が表現可能である。ベロシティでは、各トラックの音ごとに音量を 128 段階で設定することができる。本研究では、演奏表情データベース CrestMusePEDB[2]を入力データとして、メロディおよび和音進行におけるベロシティのルールを獲得する。CrestMusePEDB には、ピアノ名演奏 241

曲のスコアデータと演奏データが含まれている。MIDI ファイルの生成手法を以下に示す。

- ① 学習用の MIDI データより、発音開始時刻、ベロシティ、発音終了時刻、ノート番号の情報を取り出し、メロディおよび和音進行に分割する。
- ② 各データの演奏全体の音量が異なるため、すべてのベロシティを式(1)によりベロシティ差分に変換する。

$$vdiff(d_i, j) = v(d_i, j) - mode(d_i) \quad (1)$$

ここで、 $vdiff(d_i, j)$ は学習データ d_i の j 番目の音のベロシティ差分、 $v(d_i, j)$ は学習データ d_i の j 番目の音のベロシティ、 $mode(d_i)$ は学習データ d_i におけるベロシティの最頻値である。

- ③ メロディおよび和音進行より、それぞれ以下の4つを条件部、ベロシティ差分を帰結部としてルールを生成する。

1) メロディ

条件1：前後3音の音高の上下

条件2：前に発音された音との音高差

条件3：発音される音の音価

条件4：前に発音された音との音価差

2) 和音進行

条件1：前後3音のルート音の上下

条件2：前に発音されたルート音との音高差

条件3：発音されるルート音の音価

条件4：前に発音されたルート音との音価差

- ④ 打鍵の強弱を反映する MIDI データのメロディおよび和音進行の各音に対し、③における条件と一致するルールがあるか確認する。一致するルールがあった場合、基準ベロシティとルール

の帰結部のベロシティ差分の和を当該音のベロシティとする。一致するルールがなかった場合、当該音のベロシティは基準ベロシティとする。

3. 評価実験

本学学生 11 名、OB1 名を被験者としてアンケート形式による評価実験を実施した。3 つの楽曲に関して強弱を付けていない音源、ランダムに強弱を付けた音源、提案手法により強弱を付けた音源を用意し、被験者に聴かせる。被験者は各音源に対し、「音量の変化を感じたか」「強弱の流れは自然であったか」に関して 5 段階尺度で評価し、3 つの音源の中でどの音源が一番人間による演奏に近いかを選択して、選択理由を自由記述で回答した。評価に使用した楽曲の名称とテンポを表 1、アンケート結果の評価値の平均と標準偏差を表 2 に示す。なお、評価実験では基準ベロシティを 80 と設定する。

4. 考察

評価実験の結果から、提案手法による音源では、被験者はあまり音量の変化を感じられていないといえる。一方、ランダムにベロシティを変化させた音源では、平均 3.92 から 4.46 と他の音源と比べて高い評価値を付けていることから、多くの被験者が音量の変化を感じ取っていると考えられる。提案手法において、3 曲目の楽曲では音量の変化、強弱の流れともに中央値の 3 を超える評価が得られていることから、提案手法はテンポの速い曲に有効であるといえる。しかし 3 曲目では、どの音源が人間による演奏に近いか質問したところ、ランダムに強弱を付けた音源における評価値が提案手法を上回った。回答理由として、楽曲の強弱が大きかったことが多数指摘された。2 曲目のランダムにベロシティ変化を付けた音源では、音量の変化に関して最高値をとっているが、強弱の流れは他の音源より評価値が低かった。自由記述では、「違和感があった」、「フワフワした感じがあった」

表 1 曲名とテンポ

	曲名	テンポ
1 曲目	Piano Sonata No. 8, Op. 13, 1st Mov.	60
2 曲目	Etude No. 3, Op. 10-3	45
3 曲目	Piano Sonata K.545 1st Mov.	120

表 2 アンケート結果の評価値の平均と標準偏差

		音量の変化		強弱の流れ	
		平均	SD	平均	SD
1 曲目	強弱なし	3.54	1.01	3.77	0.80
	ランダム	3.92	1.14	3.54	1.34
	提案手法	2.92	1.33	3.31	1.26
2 曲目	強弱なし	3.08	1.27	3.31	0.91
	ランダム	4.69	0.46	2.92	1.21
	提案手法	2.85	1.10	3.08	1.00
3 曲目	強弱なし	1.69	0.82	2.92	1.33
	ランダム	4.46	0.63	3.31	1.38
	提案手法	3.69	0.82	3.85	1.10

といった意見が挙がった。1 曲目の評価では、提案手法による音源は、他の手法による音源より評価値が低かった。1 曲目の楽曲には、1 小節当たりの音の数が多く、類似したフレーズが繰り返される特徴がある。現状では、ルール抽出の際、各フレーズの繰り返し表現に関して考慮していない。フレーズの繰り返し部分では、音の強弱が単調になることが考えられる。今後の研究では、類似しているフレーズの繰り返しを考慮したルールの抽出方法を検討する。

参考文献

- [1] N.Otani, S.Shirakawa, M.Numao, "Design of Populations in Symbiotic Evolution to Generate Chord Progression in Consideration of the Entire Music Structure," Principles and Practice of Multi-Agent Systems, Lecture Notes in Computer Science, Vol.9935, pp.143-154, 2016.
- [2] 橋田光代, 松井淑恵, 北原鉄朗, 片寄晴弘, "ピアノ名演奏の演奏表現情報と音楽構造情報を対象とした音楽演奏表情データベース CrestMusePEDB の構築", 情報処理学会論文誌, Vol.50, No.3, pp.1090-1099, 2009.