

反復度に基づく特徴旋律抽出手法の提案

A Method of Characterful-Phrase Extraction Based on Adaptation

川村 修[†] 大園 忠親[†] 伊藤 孝行[†] 新谷 虎松[†]
Osamu Kawamura Tadachika Ozono Takayuki Ito Toramatsu Shintani

1. はじめに

演奏家のモデルを構築する際に用いられるものに旋律がある。既存の研究では、旋律同士の関係を学習させたり、旋律そのものからデータベースを構築することにより演奏家のモデルを構築していた。また、モデルの応用目的に対して、旋律は適切に処理される必要がある。これまでの研究では、人手によりこれらの旋律を楽曲から切り出していた。しかし、膨大な量の旋律を人手により切り出すのは、コストと一意性の点で問題がある。本研究では、スタンダード MIDI ファイルとして記録された演奏から、演奏家の特徴を示す旋律（特徴旋律）を反復度に基づいて抽出する手法を提案する。ただし、本研究で扱う旋律とは同時に2つ以上の音を発音しない単音旋律とする。本研究は、統計的な手法である反復度に基づいて自動で特徴旋律を切り出すことを試みる点で新しい。また、入力されたデータに対して出力される特徴旋律が一意に決まるので、曖昧さがなくなる点で有用である。

本稿の構成は以下の通りである。2章で関連研究について議論し、3章で反復度に基づく特徴旋律について述べる。4章で実験を行い、最後に5章で本稿をまとめる。

2. 関連研究

近年の音楽構造分析に関する研究は GTTM に基づくものが主流となっている。音楽知識表現手法に関する研究 [1] では、GTTM に基づき楽曲構造をオブジェクトとして扱うことにより、楽曲の理論的表現と基本演算を実現している。この研究は楽曲構造を自動で分析する研究として進んでいるが、旋律の自動抽出までは議論が及んでいない。

即興演奏から発音時刻の揺らぎを除去する研究 [2] では、即興演奏の発音時刻の揺らぎを確率モデルに基づき、楽譜上の正規化された位置へクオンタイズしている。この研究は旋律データベースを自動構築する一環として行われている。

本研究では、演奏中にたびたび出現する旋律とそうでない旋律との関係に着目し、反復度を用いることで特徴旋律の自動抽出を試みる。本手法では、旋律の切り出し方を式により求めることで、曖昧さをのぞくことができる。

3. 反復度に基づく特徴旋律

本研究では、同時に2つ以上の音を発音しない単音旋律を扱う。そして、旋律は曲を音列にとらえたときの部分音列とする。まず、高い頻度で出現する旋律について考える。本稿では、この旋律を一般的な旋律と呼ぶ。本研究であつかう特徴旋律は、一般的な旋律と一般的な

な旋律の組み合わせによる旋律とする。ここでいう特徴旋律は演奏家に特有な旋律の展開方法を示していると考えられる。以下では、特徴旋律を求めるための定義を行う。

3.1 音列と音

曲 s 中に含まれる n 個の音列 p を $p = (p_1, p_2, \dots, p_i, \dots, p_n)$ とする。 p_i は1つの音符（休符）を表す。ただし、 s が休符で始まる時最初の休符は無視し、 p_1 は音符を示すこととする。 p_i の音高を表すため $pitch(p_i)$ は、 p_i が音符の場合 MIDI ノートナンバー[‡]を与え、 p_i が休符の場合 $pitch(p_{i-1})$ を与えるとする。 p_i の音長は、 $gatetime(p_i)$ で示す。ただし、 p_i が4分音符（休符）のとき $gatetime(p_i) = 480$ とし、全ての p_i において $gatetime(p_i) > 0$ とする。

ある2つの旋律が全く等しい音高の列となっていることは少ない。したがって、旋律の出現頻度を求める際、音高の定義がこのままでは得られる出現頻度は極端に低いものになってしまう。そこで、音の高さを音高ではなく前の音高との差分を半音の数で表すことにする。この差分を式 (1) で定義する。ただし、 $degree(p_1) = 0$ とする。

$$degree(p_i) = pitch(p_i) - pitch(p_{i-1}) \quad (1)$$

3.2 旋律の比較と反復度

曲を構成する音列の部分音列全てについての出現頻度を求めるのは、膨大な計算量とメモリを必要とする。1曲を構成する音の数を k とすると、部分音列の数は $k(k+1)/2$ となる。最悪の場合それらすべてのマッチングを行わなければならない。したがって、マッチングのオーダは $O(N^2)$ となる。具体的には1曲450個の音があるとすると、1曲あたりの部分音列は 10^5 個程度になり、10曲を扱うとするとマッチング回数は最悪 10^{12} 回程度となる。また、すべての部分音列に対して出現頻度を保存する場合、 $O(N^2)$ の表に相当するメモリが必要となる。

本研究では、Suffix Array [3] を用いて計算量を $O(N \log N)$ およびメモリを $O(N)$ にする。部分音列に対して順序付けを行うために、一つの音を一つの整数で示す。音列の出現頻度を測る際、音列を構成する各音の比較を行う必要があるが、音の比較には、音の長さや高さを比較する必要がある。旋律 q を $q = (q_1, q_2, \dots, q_n)$ とする。 q_i は一つの音を示し、式 (2) で定義する整数値をとる。ここで $sign(p_i)$ は、 p_i が音符のとき $degree(p_i) \geq 0$ ならば1、 $degree(p_i) < 0$ ならば-1となり、 p_i が休符のとき-1となる値をとる。2つの旋律が等しいとは、各々の旋律を構成する音の数が等しく、かつ、対応する音の整数値が等しいことである。

$$q_i = sign(p_i) \times \{10^3 \times gatetime(p_i) + |degree(p_i)|\} (2)$$

次に、一般的な旋律を求めるために音声処理や言語処理に用いられている反復度 [4] を用いる。ある旋律が

[†]名古屋工業大学 大学院工学研究科, Graduate School of Engineering, Nagoya Institute of Technology

[‡]各音に0以上127以下の整数を一意に割り当てたもの

表 1: Miles Davis の特徴旋律の一部

一般的な旋律	特徴旋律 (phrases)
70000,-10000,-70002	70000,-10000,-70002,-10000,-1295001,-145000,860001,-100000,645002,-75000,-215002,-25000,-1295001,-625000,-140007
215000,-25000,-215004	215000,-25000,-215004,-505000,430013
215000,-25000,-215004	215000,-25000,-215004,-25000,-320005,-40000,-105004,-15000,430008

様々な曲に共通して出現するだけでは、一般的な旋律と見なすのは適切でない。一般的な旋律が用いられる曲では、その一般的な旋律は繰り返し用いられるはずである。したがって、本論文では反復度を用いることとした。本研究での反復度とは、ある曲 s に q が 1 回以上出現することを条件としたときに、 s に q が 2 回以上出現する条件付き確率である。曲集合 $S = (s_1, s_2, \dots, s_n)$ があるとき、 S 全体で q を含む曲の数を $df(q)$ 、 S 全体で q を 2 回以上含む曲の数を $df_2(q)$ とすると、 q の反復度は $adaptation(q) = df_2(q)/df(q)$ と定義される。以上から特徴旋律の集合 *phrases* を式 3 により定義する。ここで、 $head(q) = \max_{t \in T} \{adaptation(t)\}$ 、 $tail(q) = \max_{o \in O} \{adaptation(o)\}$ とする。ただし、 T は q_1 から始まる q の部分列集合、 O は q_1 を除いた q の部分列集合を示す。また、 α はしきい値を示す。

$$phrases = \bigcup_q \{ head(q) \geq \alpha \wedge tail(q) < \alpha \wedge adaptation(q) < \alpha \} \quad (3)$$

4. 特徴旋律抽出実験

4.1 実験

反復度を用いて、演奏家の演奏データから特徴旋律の抽出実験を行う。ジャズトランペット奏者の Miles Davis と、ジャズピアニストの藤井英一氏の 2 名についてそれぞれ実験を行った。演奏データの元となる曲として、Miles Davis では 19 曲 [5]、藤井英一氏では 5 曲 [6] をそれぞれ用いた。これらの曲を、音符分解能が 4 分音符を 480 として入力した。しきい値は予備実験により、 $\alpha = 0.7$ とした。実験により抽出された特徴旋律の一部を Miles Davis については表 1 に、藤井英一氏については表 2 に示す。

4.2 考察

しきい値について考察すると、値 α を下げることにより、抽出される特徴旋律は一般的な旋律に近い旋律となる。一方、 α を上げることにより、より長い旋律を抽出することができる。

2 人の演奏家から抽出された特徴旋律の全体を比べてみると、Miles Davis は比較的長い旋律が、藤井氏では比較的短い旋律が多かった。これは違う演奏家に同じしきい値 ($\alpha = 0.7$) を用いたことを考えると、Miles Davis の

表 2: 藤井英一氏の特徴旋律の一部

一般的な旋律	特徴旋律 (phrases)
215000,-25000,-215004	215000,-25000,-215004,-25000,105001
215000,-25000,-215004	215000,-25000,-215004,-25000,-430004,-530000,215007,-25000,-43002

旋律は単純な演奏の組み合わせではなく、複雑な演奏の組み合わせから成り立っており、同様の演奏を行うことは比較的少ないといえる。藤井氏の旋律が比較的短いのは、短い音の連続で似たような演奏をたびたび行っていたことにより、 $tail(q)$ がしきい値を超えることが多かったからだと考える。実験により抽出された旋律には、ある一般的な旋律から始まる数パターンの特徴旋律が含まれていた。このことから、本研究により得られた特徴旋律は、演奏家の旋律のバリエーションという特徴を示していると考えられる。また、表 1 と表 2 からわかるように、違う演奏者同士でも、同じ一般的な旋律を含む特徴旋律がある。しかし、一般的な旋律に続く一般的でない旋律が異なっている。このことから、本手法により演奏者に特有な旋律の展開方法を反映した旋律、つまり、旋律のバリエーションを抽出できているといえる。

5. まとめ

本研究では、反復度に基づく特徴旋律の抽出手法の提案を行った。これまで、旋律データベースに格納する旋律を曲から切り取る作業は、人手により行われていた。本研究は、統計的な手法である反復度に基づいて自動で特徴旋律を切り出す点で新しい。本手法では、反復度を用いることにより一般的な旋律を求め、その旋律から派生する様々な旋律を抽出することで演奏家の旋律のバリエーションという特徴を抽出した。

参考文献

- [1] 平田圭二, 青柳龍也: 音楽理論 GTTM に基づく多声音楽の表現手法と基本演算, 情報処理学会論文誌, Vol.43, No.2, pp.277-286(2002).
- [2] 浜中雅俊, 後藤真孝, 麻生英樹, 大津展之: 発音時刻の楽譜上の位置を確率モデルにより推定するクオンタイズ手法, 情報処理学会論文誌, Vol.43, No.2, pp.234-244(2002).
- [3] 梅村恭司, 真田亜希子: 文字列を k 回以上含む文書数の計数アルゴリズム, 言語処理学会自然言語処理, Vol.9, No.5, pp.43-70(2002).
- [4] 武田善行, 梅村恭司: キーワード抽出を実現する文書頻度分析, 計量国語学, Vol.23, No.2, pp.1-26 (2001).
- [5] Miles Davis: Standards Volume 1, HAL LEONARD Corp.
- [6] 藤井英一: 実例ジャズ・ピアノ・アドリブ集 2, リズム・エコーズ (1999).