

武藤 誠 半田 伊吹 日比 啓文 坂井 修一 田中 英彦

{muto, handa, hibi, sakai, tanaka}@mtl.t.u-tokyo.ac.jp

東京大学大学院 工学系研究科 *

1 はじめに

音楽は音色、旋律、リズム、和声などの複数の構成要素から成り立っている。これらの構成要素の主観的な類似度を計算することによって、感性情報抽出などの工学的应用が可能となる。本稿では音楽構成要素の1つである旋律輪郭・音程の主観的類似度¹の計算手法を提案する。従来手法では音高の要素としてピッチのみを考慮しているため、オクターブ類似性²などの音楽の知覚特性を計算結果に反映できないという問題があった。そこで本稿では、ピッチ、半音音度、5度圏を考慮した心理的音高モデルに基づいた主観的類似計算の手法を提案する。評価実験の結果、従来手法と比べてより正確に類似度計算を行なうことができた。

2 音楽構成要素の主観的類似度計算

2.1 音楽構成要素

音楽は旋律、リズムなどの複数の構成要素から成り立っており、我々はそれらの構成要素を並列的に知覚する [1,2]。本研究では音楽構成要素として、図1のように楽曲の各パート毎に、音色、旋律、リズム、和声を属性とするものを想定する。








	音色	旋律	リズム	和声
旋律パート	楽器音 	旋律輪郭・音程 -4,+7,-3,0,...	リズムパターン 	音名頻度 
伴奏パート	楽器音 	伴奏パターン 	リズムパターン 	コード 
...

図1: 音楽の構成要素

*"Computing the Psychological Similarity of Musical Components"

Makoto Muto, Ibuki Handa, Hirofumi Hibi, Shuichi Sakai, Hidehiko Tanaka

University of Tokyo, Graduate School of Engineering,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

¹物理的類似度に対し、心理的な類似度を主観的類似度とする。

²基音が整数倍の関係にある2つの音が類似した印象を与える現象

2.2 各音楽構成要素の主観的類似度計算

各音楽構成要素の主観的類似度を計算し、非類似度を多次元空間上の距離で表現することによって、図2のような主観的類似度マップ³が得られる [3]。この類似度マップによって次項のような様々な応用技術が可能となる。

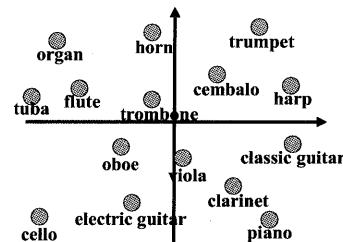


図2: 音色の主観的類似度マップ

2.3 応用分野

感性情報抽出

感性情報抽出の従来手法では、楽曲の特微量と感性情報との相関分析⁴が行なわれてきたが、この手法は感性情報抽出には適さない [2]。そこで、本研究では事例ベースの手法を感性情報抽出に適用することを考えている。事例ベースシステムでは、分析対象の類似度関係の計算が必要であるが、音楽はその多様性ゆえに類似度計算が困難である。しかし、各音楽構成要素ごとの類似度計算は比較的容易であるので、各構成要素の類似度から全体の楽曲の類似度の推定が可能となる。これによって、事例ベースによる感性情報抽出が可能となる。

音楽のモーフィング

各音楽構成要素の類似度マップ上で連続的に変化させた構成要素を用いて楽曲断片を合成し、順次並べて演奏することによって音楽のモーフィングが可能となる。

主観的類似度に基づく楽曲検索

各構成要素の類似度から全体の楽曲の類似度の推定することによって、楽曲の類似度マップを作ることができる。これによって、主観的に類似した楽曲を検索する技術が可能となる。

³図では2次元の場合を示したが、より多くの次元数となり得る。

⁴重回帰分析やニューラルネットワーク機械学習など

3 旋律輪郭・音程の主観的類似度計算

3.1 関連研究

従来手法では、音楽検索における旋律の類似度計算への応用を想定して、旋律間の類似度をDPマッチングによって計算することが行なわれている[4]。これらの手法では旋律の音高の要素としてピッチのみを想定している。しかしながら、この手法ではオクターブ類似性など、主観的類似度に影響を与える音楽の知覚特性を扱うことができない。

3.2 提案手法

本研究では心理的な音高の多次元空間モデルを設定し、旋律の輪郭・音程を多次元空間上の時間変化として表現する。そして、旋律間の主観的類似度を多次元空間上のベクトル表現を用いて定式化する。

音高の心理的モデル

音高の心理的モデルは様々なものが提案されている[5]。本研究では多くのモデルで共通して考慮されている、ピッチ、半音音度、5度圏の3つの要素による音高モデルを設定する。

ピッチはMIDIのノートナンバー a で表す。半音音度、5度圏はそれぞれ図3のように2次元座標上の円上の点 $b = (x_1, y_1)$, $c = (x_2, y_2)$ として表す。これらを用い、心理的音高を $\{a, b, c\}$ で表す。

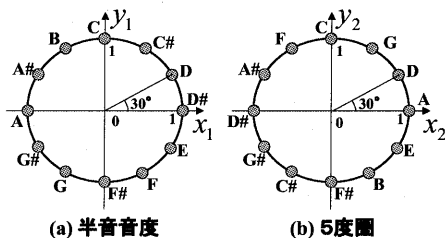


図3: 半音音度と5度圏の座標表現

類似度計算

本稿では比較対象の旋律の音符数は等しいとする。旋律の輪郭・音程を $\{a, b, c\}$ の時系列 $\{a_t, b_t, c_t\}$ で表す。各 t はそれぞれ1つの音符に相当する。 $\{a_t, b_t, c_t\}$ の時間差分を $a'_t = a_{t+1} - a_t$ などとする。2つの輪郭・音程 $\{a'_1, \dots\}$, $\{a'_2, \dots\}$ の非類似度を $\sum_t \{w_1 |a'_1 - a'_2| + w_2 |b'_1 - b'_2| + w_3 |c'_1 - c'_2|\}$ とする。ここで、 w_n は音高モデルにおける各要素の主観的類似度に与える影響の度合の重みを表す⁵。



図4: 実験に用いる曲の例

3.3 評価実験

方法

図4のように、ハ長調におけるC4からB4までを1回ずつ用いて8分音符で並べた旋律を10曲用意した。音色は正弦波とし、テンポは120とした。このような制約により、本研究が想定する音色、リズム、和声属性を一定とみなすことができ、旋律属性について選択的に実験を行なうことができる。これらの曲を図4のように全ての2つの順列の10P2 = 90曲を用意した。これらをランダムな順序で被験者1人に提示し、非類似度を0から10の11段階で評価した。次に、同一曲の提示順序の違う2曲に対する評価を平均し、主観的非類似度とした。

一方、用意した曲について前項のように類似度計算を行なう。結果の非類似度は順序尺度の性質を持つと考えられるので、計算結果と被験者による評価との相対的な大小関係を評価対象とする。具体的には、重複を除いた45曲の全組み合わせ45C2 = 990通りにおける、計算結果と被験者の評価の大小関係の一致する割合を求めた⁶。

結果・考察

類似度計算の重み w_n はノートナンバー、半音音度、5度圏の影響を等しく設定するために、各要素の期待値で正規化し、 $w = \{1, 3.28, 4.03\}$ とした。その結果、922個中548正解の59%の精度であった。比較のため、 $w = \{1, 0, 0\}$ として、半音音度、5度圏の影響を無しとした場合、902個中538正解の58%であり、正解率が下がった。この結果から、旋律の主観的類似度計算における、半音音度、5度圏を考慮することの有効性が予想される。

4 まとめ

本稿では音楽が複数の構成要素から成り立っており、それぞれの構成要素の主観的類似度の計算により様々な工学的応用が可能となることを述べた。次に、音楽構成要素の1つである旋律輪郭・音程の主観的類似度の計算手法としてノートナンバー、半音音度、5度圏を考慮したものを提案し、評価実験によってその有効性を示した。

参考文献

- [1] 岡田顕宏, 阿部純一: 「音楽知覚におけるモジュール性」, 音楽知覚認知研究, Vol.5, No.1, pp.22-43, 1999.
- [2] 武藤誠, 木下智義, 半田伊吹, 坂井修一, 田中英彦: 「音楽聴取時における感性の計算モデルの提案」, 情報処理学会第60回全国大会講演論文集(2), pp.51-52, 2000.
- [3] 鈴木健嗣, 橋本周司: 「ニューラルネットワークを用いた感性情報の数量化」, 電子情報通信学会論文誌, Vol.J82-D-II, No.4, pp.677-684, 1999.
- [4] 柳瀬隆史, 高須淳宏, 安達淳: 「音楽検索における自動インデクシング法」, 情報処理学会研究報告, 98-DBS-116(2), pp.117-124, 1998.
- [5] Fred Lerdahl: Tonal Pitch Space, *Music Perception*, Vol.5, No.3, pp.315-350, 1988.

⁵ w_n の値は音楽の聞き方の個人差に応じて設定する必要がある。

⁶ 被験者の類似度評価の等しい曲の組み合わせについては評価対象から除外した。