

形容詞対を用いた音楽データベース検索システムに関する研究(Ⅱ)

5G-4

池添 剛 梶川 嘉延 野村 康雄

関西大学工学部電子工学科

1. はじめに

近年、インターネットや MIDI (Musical Instrument Digital Interface) などの普及により、デジタル情報として蓄積される音楽情報が急増しており、それらの検索システムに対する要求が高まっている。その一つの手法に、感性語を用いた検索システムがある。我々は、従来の検索システムの問題点を解決するために、新たな検索システムを提案し、先に報告した^[1]。しかし、その検索システムでは、試聴実験を行った曲しか検索を行う事ができなかった。その問題点を解決する一手法として、自動インデクシングシステムの導入が考えられる。しかし、従来の自動インデクシングシステム^{[2][3]}にはいくつかの問題点がある。そこで本稿では、新たな自動インデクシングシステムの提案を行い、その構築結果を示す。

2. 形容詞対を用いた音楽データベース

検索システムの問題点

文献[1]にも述べたように、我々が提案してきた形容詞対を用いた音楽データベース検索システム（以下、検索システム）では、試聴実験結果を元にして検索空間となる音楽感性空間を生成している。そのため、検索システムでは試聴実験を行っていない曲を検索する事ができないという問題点があった。しかし、試聴実験を行っていない曲の曲特徴量（曲の特徴を数値化したもの）から音楽感性空間座標値を求めることができれば、その問題点は解決できる。そこで、その問題点を解決する一手法として、試聴実験を行っていない曲の音楽感性空間座標値を求める、自動インデクシングシステムの提案を行う。

従来の自動インデクシングシステムに関する研究

A Study on Music Database Retrieving System by Pair of Adjective (Ⅱ).

Takeshi Ikezoe, Yoshinobu Kajikawa,

Yasuo Nomura

Department of Electronics, Faculty of Engineering,

Kansai University

3-3-35, Yamate-cho, Suita-shi, Osaka, 564-8680, Japan

には、佐藤ら^[2]の研究や、柳瀬ら^[3]の研究などがある。しかし、柳瀬らのシステムは、楽曲の主旋律中のフレーズを類似度に基づいてグループ分けし、それぞれにラベル付けを行うというものである。本研究の目的とは異なる。また、佐藤らの研究では、曲特徴量と曲印象は線形的な関係という仮定の元で、重回帰分析を用いてシステムの構築を行っている。しかし、著者らも述べているように、曲特徴量と曲印象が線形的な関係ではない可能性がある。

そこで、本稿で新たな自動インデクシングシステムの提案を行い、検索システムの問題点を解決する。

3. 自動インデクシングシステム

図1に今回提案する自動インデクシングシステムの概要を示す。本システムの特徴は、遺伝的アルゴリズム（以下、GA）を用いて曲印象に関わりのある曲特徴量の選出を行い、それをを用いて重回帰分析を行う点である。また、GAを用いる際に、曲特徴量同士の積も選出されるようにする事で、重回帰分析で非線形的な予測を可能としている。

ここで、図1について説明する。まず、MIDIデータから各音符の音程や強さといった演奏情報を抽出し、それをを用いて表1に示す曲特徴量を計算する。次に、GAを用いてそれらの曲特徴量から曲印象に関わりのある曲特徴量を選出する。なお、GAの適

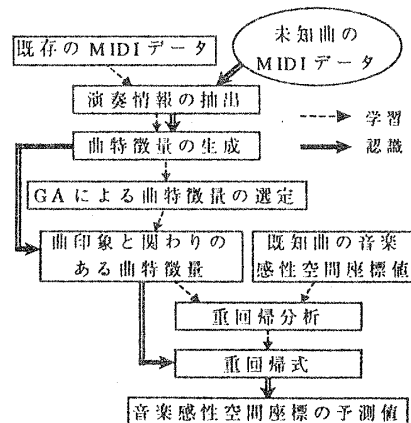


図1 自動インデクシングシステムの構成図

応度には、重回帰分析の予測精度を表すパラメータの一つである、自由度調整済み重相関係数を用いた。そして、選出された曲特徴量を説明変数、実際の音楽感性空間座標値（以下、実際値）を目的変数として重回帰分析を行い、重回帰式を求める。このように求めた重回帰式の説明変数に未知曲の曲特徴量を代入することで、未知曲の音楽感性空間座標の予測値（以下、予測値）が求まる。

表1 曲特徴量

| 曲特徴量 | 意味 |
|-------------------------------|-----------------------|
| Ave(S)-Pitch* | 音の高さの平均値(分散値) |
| Ave(S)-Velocity* | 音の強さの平均値(分散値) |
| Ave(S)-Duration | 音の長さの平均値(分散値) |
| Rest | 音の鳴っていない時間の割合 |
| Density | 音数の時間平均値 |
| Differential | 平均音高からの絶対値の平均 |
| Resister | 音高の最大値と最小値の差 |
| Harmony** | 音の調和度 |
| Ave(S)-Rhythm** | リズムの平均値(分散値) |
| Key** | 曲の調性 |
| Meter** | 曲の拍子 |
| Tempo** | 曲のテンポの初期設定値 |
| Low(Middle)[High] Note Number | 低音(中音)[高音]の音数の割合 |
| Low(Middle)[High] Note Time | 低音(中音)[高音]が鳴っている時間の割合 |

* : 音の長さを重みとして掛け合わせて計算される

** : MIDIのメタデータからの情報を用いて計算される

4. 実験結果

ここでは、3.で述べた自動インデクシングシステムを用いて未知曲の音楽感性空間座標値を生成する。なお、実験には、常盤女子短期大学苦瓜瑞恵女史演奏による「ブルグミュラー25の練習曲」の25曲を用いた。表2にその条件でインデクシングを行った結果を示す。なお表2は、25曲の内、特定の1曲を除いて重回帰式を求め、その式で特定の1曲の予測値を求めるといった方法を25曲に対して行った結果の値である。表2を見ると、非常に高い精度でインデクシングできている事が確認できる。

次に、この予測値を用いて、実際に検索システムで検索を行う。具体的には、ある特定の1曲は予測値、それ以外は実際値を用いて音楽感性空間を作成し、試聴実験結果から得られた特定の1曲に対する形容詞対の値を検索システムに入力した時に、特定の1曲が検索されれば正解とする方法を25曲に対して行う。その結果を表3に示す。なお、中間層のユニット数7-7とは、ニューラルネットワーク（以下、

NN)を用いた検索システムにおけるNNの中間層の層数が2層であり、それぞれ20ユニットからなる事を表す。また、不正解曲中の網掛け部分は、実際値を用いた場合でも不正解であった曲である。さらに、表3に示す三つの規模のNNを利用した理由は、この三つが一番高い検索精度を得ていた^[1]ためである。ここで不正解であった曲を考察すると、全体的な傾向としては、11曲目と25曲目を誤認識していた。しかし、これらの音楽感性空間座標値は非常に近い距離にあるため、曲印象の近いところで誤認識していると言える。また、中間層のユニット数7-7では15曲目を20曲目と検索しているが、これは実際値を用いた場合でも誤認識を起こしていた曲であるため、自動インデクシングシステムが誤認識の原因ではない。むしろ、実際値を用いた場合と同じ曲を誤認識していることより、自動インデクシングシステムの精度は非常に良いと言える。

表2 インデクシング結果

| 相関係数 | 平均二乗誤差 | 予測値と実際値の最大平方距離 |
|--------|------------------------|-------------------------------|
| 0.9992 | 4.562×10^{-4} | 1.2606×10^{-2} (7曲目) |

表3 検索結果

| 中間層のユニット数 | 正解数 | 正解率 [%] | 不正解曲 |
|-----------|-------|---------|------------------------------|
| 20-20 | 21/25 | 84 | 11→25, 15→20 22→19, 25→11 |
| 10 | 22/25 | 88 | 3→8, 5→17 11→25 |
| 7-7 | 23/25 | 92 | 11→25, 15→20 |

5. まとめと今後の方針

本稿では、我々が提案してきた検索システムの問題点の一つを解決するために、新たな自動インデクシングシステムの提案を行い、その有効性を示した。その結果、試聴実験を行っていない曲に対しても検索を行う事が可能となり、問題点が解決された。

しかしながら、検索システムには特定の形容詞対による検索ができないという問題点が残っている。そこで、今後はその問題点を解決していきたい。

【参考文献】

- (1)池添他:第58回情処全大,Vol2,pp.109-110,Mar.1997.
- (2)佐藤他:情処研報 DS118-8,pp.57-64,May.1999.
- (3)柳瀬他:情処研報 DS116-42,pp.117-124,Jul.1998.