

MIDI データと自己組織化マップを組み合わせた 演奏曲調分類の研究

垣田 詩鶴香 三好 力

龍谷大学理工学部

1. はじめに

一般的に人が何らかの用途で楽曲を選ぶ際は、歌手・作曲家・ジャンル・曲調等の特徴を利用していると考えられる。楽曲の曲調による分類に注目すると、たとえば、街中を歩いているとあちこちのお店から音楽が聞こえてくるが、お店の雰囲気合った曲調の音楽を流していることが多いように感じる。また、曲調分類は、音楽を聴く際に使うことも出来るが、楽器を練習するための練習曲を選ぶ際にも有用である。従来練習曲は、基本的には既知の曲や作曲家・演奏者などから選び、有名な曲や有名な作曲家など、偏った選曲になることが多かった。しかし曲調分類を行うことが可能になると、知らない曲からも選ぶことが出来、より好みの曲や作曲家を見つけることの出来る機会が増える。また、演奏した情報を入力に用いると演奏曲と同じような曲調の曲や異なる曲調の曲にどのような曲があるかを簡単に調べることが出来る。練習曲に同じような曲調の曲を選べばより技術力や表現力を深めることが出来るし、異なる曲調の曲を選べば技術や表現の幅を広げることが出来ると考えられる。このように曲調は選曲の上で重要度が高いと考えられるが、歌手や作曲者で選別することに比べて、ジャンルや曲調は感性による分類であるため、それらで選別するのは容易ではない。そこで本研究では自己組織化マップ(Self-Organizing Map:以下 SOM とする)と MIDI 形式のデータ(以下 MIDI データとする)を用いた曲調分類を検討した。

2. SOM

SOM は教師なし学習の手法の一つである。汎化能力が高く結果もわかりやすいため、容易に曲調で選別するために適している。SOM を用いた音楽分類の他の研究である「自己組織化マップを用いた MIDI 音楽の分類」[2]では音高を入力としているため、それを参考に本研究では音高と音の長さを用いて曲調分類を行うことを想定している。

3. MIDI

Musical Instrument Digital Interface(以下 MIDI とする)とは、電子楽器の演奏データを機器間でデジタル転送するための世界共通規格である。MIDI

データはヘッダ部とトラックデータ部からなり、トラックデータ部はデルタタイム、システムコマンド、音高、音の長さ、音の強さの各情報単位を組としたデータ列からなる。[1]情報単位は 8bit、すなわち 16 進数 2 桁からなり、デルタタイムと音高、音の長さ、音の強さは 0x00~0x7F の値を、システムコマンドは 0x80~0xFF の値をとる。MIDI データは音響波形情報(以下波形情報とする)と比べてデータ量が少なく、また、MIDI に対応した電子楽器を演奏してその演奏情報を計算機に入力するこの形式で保存されるため、処理が容易である。波形情報から音高や音の長さを得るためにはフーリエ変換などの処理が必要になるため、全ての処理を行うには多くの時間が必要になり、計算量が大きくなる。また、演奏曲の入力データも波形情報を用いるよりも MIDI 楽器の演奏情報から得た MIDI データを用いた方がフーリエ変換等の処理が不要であるため処理が容易である。これらの理由から、本研究では波形情報ではなく MIDI データを用いる。

4. ジャンルと曲調

音楽におけるジャンルや曲調とは、定義付けが曖昧である。本研究では、“クラシック”や“ロック”、“ポップス”などの音楽の様式や形式のことをジャンル、“勇ましい”、“華やか”などの楽曲の調子や雰囲気のことを曲調と定義する。

5. 実験 1

MIDI データから得られる、全ての楽器の音高、長さ、強さを総合した音の流れと、前述のジャンルの間に関係性が見られるかどうかを確認する実験を行った。MIDI データにはヘッダ部とトラックデータ部が存在するが、ヘッダはトラックデータと比べてデータ量が遥かに小さいため、無視できる程度の誤差である。また、トラックデータには音高、音の長さ、音の強さ、システムメッセージの 4 種類のデータが存在するが、音高、音の長さ、音の強さは 0x00~0x7F の範囲、システムコマンドは 0x80~0xFF の範囲で表されるため、0x00~0x7F のデータを用いることによってシステムコマンドを除去する。メロディーとジャンルの間に何かしらの関係性がある場合、それぞれのジャンルによって音高と音の長さ、音の強さの特徴があると考えられるが、音高、長さ、強さは独立した情報であるために特徴は異なり、音高と音の長さ、音の強さの分布を足し合わせて一つの分布として見ても

「Study of the melody style classification of MIDI data using a SOM」

KAKITA Shizuka, MIYOSHI Tsutomu
Faculty of Science and Technology, Ryukoku University

ジャンルごとの特徴が失われないと考えた。よって、MIDI データの音高と音の長さ、音の強さの情報を分離せずに用いて分類を行ってもある程度の性能で分類が可能であると考えた。

実験には様々なジャンルと曲調の楽曲 102 曲を用意し、それらを入力ベクトルとして SOM を作成した。出力は 10×10 ユニットで行った。なお、今回用いた楽曲は RWC 研究用音楽データベースから得た 102 曲であり、ジャンルはデータベース上に定義されている、ポップス、バラード、ロック、ヘビメタル、ヒップホップ、ハウス、テクノ、ファンク、ソウル、ビッグバンド、モダンジャズ、フュージョン、ボサノバ、サンバ、レゲエ、タンゴ、クラシック、ブラスバンド、ブルース、フォーク、カントリー、ゴスペル、アフリカン、インド、フラメンコ、シャンソン、カンツォーネ、演歌、民謡、雅楽の 30 ジャンルを用いた。

5.1. 実験 1 のデータ

MIDI データのバイナリデータから得た 16 進数 2 桁の各文字のヒストグラムを用いた。音高、音の長さ、音の強さは $0x00 \sim 0x7F$ の範囲のみで表されるため、 $0x80 \sim 0xFF$ の範囲のデータを除いた 128 次元のベクトルを作成して 1 曲分のデータとした。

5.2. 実験 1 の結果

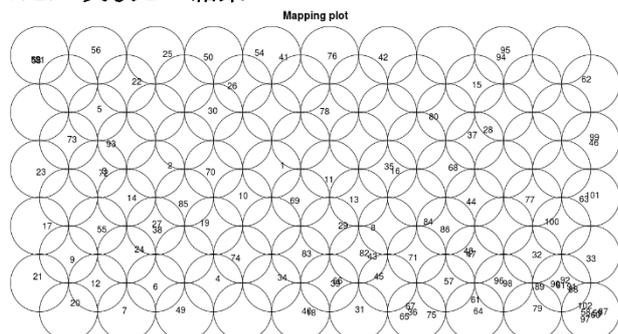


図 1: 入力 128 次元、出力 100 ユニットの出力結果
図 1 はその SOM である。図より、30 分類中 23 分類、101 楽曲中 61 楽曲が同ジャンルの他楽曲と近い位置に配置されたことが確認された。同ジャンルの楽曲と同じユニットに出力されたのは 101 楽曲中 22 楽曲である。ヒストグラム作成には約 1.128 秒の処理時間を要した。

6. 実験 2

さらなる計算量削減を考える。音高、長さ、強さはそれぞれ 7bit16 進数 2 桁で表現されるため実験 1 では 128 次元のベクトルを用いた。しかし、多くの楽曲はその一部分の値のみを用いている事に着目し、桁を無視して MIDI データを 8bit16 進数 1 桁の集合とみなして読み込み、16 次元のヒストグラムをデータベクトルとしても楽曲の特徴を保持しているのではないかと考えた。これにより、MIDI データからのヒストグラムの作成時の処理を低減することができる。さらに、128 次元データの 8 分の 1 の 16 次元データを用いて SOM の学習を行

うため、学習時間が短縮される。

6.1. 実験 2 のデータ

MIDI データのバイナリデータから得た 16 進数 1 桁の各文字のヒストグラムを用いて 16 次元のベクトルを作成して 1 曲分のデータとした。なお全ての文字を用いたため、実験 1 では除いた $0x80 \sim 0xFF$ の範囲も含まれている。

6.2. 実験 2 の結果

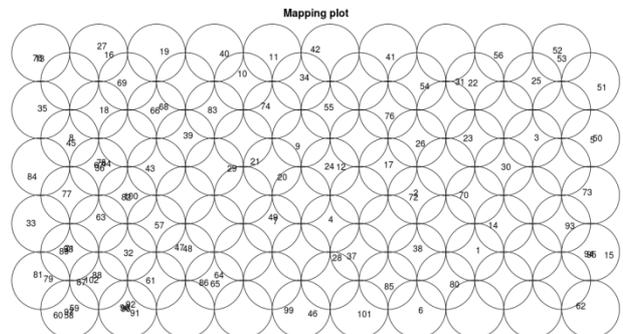


図 2: 入力 16 次元、出力 100 ユニットの出力結果
図 2 はその SOM である。図より、30 分類中 22 分類、101 楽曲中 63 楽曲が同ジャンルの他楽曲と近い位置に配置されたことが確認された。同ジャンルの楽曲と同じユニットに出力されたのは 101 楽曲中 20 楽曲である。ヒストグラム作成には約 1.058 秒の処理時間を要した。

7. 考察

実験 1 では約 60.4%、実験 2 では約 62.4%の楽曲が分類できることが確認された。ヒストグラムを作成する処理時間は 128 次元と比べて 16 次元は約 6.28%減少した。精度という観点では、同じユニットに出力された数の多い 128 次元の方が優れているが、分類数という観点では 16 次元の方が優れていると言える。これは、128 次元から 16 次元へと変化した際に音楽と関係のない $0x80 \sim 0xFF$ のデータが増加したことにより、汎用性が高くなる代わりに精度が減少したと推測される。しかし、精度においても分類数においてもほとんど差のないことを考慮すると 16 次元で扱っても問題はないと考える。処理時間が短い点や本研究での目的においては精度よりも汎用性を重視するため、今回の目的には 16 次元の方が適している。分類システムとして利用可能なレベルではないが、過半数の楽曲が分類可能ということは、MIDI データから得られるメロディーとジャンルとの間に何らかの関係性があることが示唆される。今後は簡易的な処理のまま分類精度を上げる方法を検討していきたい。

参考文献

- [1]高橋信之, コンプリート MIDI ブック, 株式会社リットーミュージック, 2005 年
- [2]田中耕平, 堂蘭浩, “隠れマルコフモデルと自己組織化マップを用いた MIDI 音楽の分類”, 電子情報通信学会技術研究報告. NC, ニューロコンピューティング 108(281), 43-48(2008)