

関数データ解析を用いた楽曲からの特徴抽出

函館高専情報工学科 東海林智也*

1 はじめに

インターネット上での楽曲配信サービスが拡大するにつれて、大量の楽曲の中からある与えられた条件を満たす特定の楽曲を検索する技術が重要となってきた。楽曲に含まれるいくつかの特徴量を抽出することによって楽曲データベースを作成し、そのデータベースを元にしてこれらの検索はこなわれている [1]。

従来の楽曲からの特徴抽出手法としては、旋律に含まれるいくつかの特定パターンの出現確率を調べる手法 [2]、隠れマルコフモデルを用いて旋律の遷移状態を調べる手法 [3] などの局所的な特徴量を対象とした抽出手法と、旋律に含まれる音符の高さや長さの割合を取り出す手法などの大域的な特徴量を対象とした抽出手法の 2 つの手法に大別できる。しかし局所的な特徴量を対象とした抽出手法も大域的な特徴量を対象とした抽出手法も、旋律が持つ連続的なデータとしての特徴を十分に生かしているとは言えない。

そこで本稿では楽曲の旋律を連続的な関数データとみなして関数データ解析 [4][5] を適用することによって楽曲の持つ特徴を抽出することを試みる。サンプリング時刻に関して自由度の高い関数データは局所的な特徴量も大域的な特徴量も含む中間的なデータとみなすことができるため、旋律が本質的に持つ特徴をより際立たせることが可能であると期待できる。なお本稿ではレジストレーションの問題を避けるために、旋律を時間領域での関数データとみなしてそのまま関数データ解析を適用するのではなく、フーリエ変換を用いて作成した 2 組の周波数領域における関数データの特徴量として利用している。さらに適用例として、個人の楽曲に対する嗜好性 (好き・嫌い) 情報を関数データを用いた判別分析を用いて抽出する例を示す。

2 関数データ解析

Ramsay らにより提唱された関数データ解析はデータを離散値としてではなく連続な関数として扱うデータ解析手法である [4][5]。離散的なデータを補間を用いて連続関数として表すことによってサンプリング時刻に関する自由度が上がるため、等間隔時刻でサンプリングをおこなう必要性が無くなるという利点がある。また一般的な主成分分析、回帰分析、クラスター分析、判別分析などの手法も関数データ解析に拡張されて用いられている。

3 旋律からの関数データの作成

旋律上の音符を任意の間隔時刻でサンプリングした離散的なデータ列とみなせば、五線譜はそのまま横軸が時間、縦軸が音の高さを表す点プロット図となる。従って前節で述べた関数データ解析の概

*tokai@hakodate-ct.ac.jp

念を楽曲に対してそのまま適用することが可能である．本稿では旋律から次のようにして関数データを作成する．なお旋律として楽曲の特徴を最も多く含むと思われる主旋律を用いることにする．

Step 1. 楽曲の主旋律全体から特徴を抽出したい部分を選択して切り出す

Step 2. 各音符の音の高さに関して適当な音程番号 (例えば中心周波数や MIDI フォーマットのノート番号など) を割り当てる

Step 3. Step 2 で割り当てた番号と音符が鳴り始める時間から点プロット図を作成する．このプロット図の横軸は時間，縦軸は音程番号を表している

Step 4. プロットされた点の間を直線または曲線で補間する

この時点で旋律上の音符から直線や曲線の補間により作られた時間領域における関数データが得られるが，標本となる各楽曲間で演奏時間が異なるため時間領域での関数データをそのまま使用して関数データ解析をおこなう場合には時間軸に関してレジストレーション (伸縮による調整) をおこなう必要がある．しかし主旋律は関数データの形状，すなわち時間軸自体にも意味があるため，機械的にレジストレーションをおこなうと主旋律に含まれる特徴成分が破壊される恐れがある．従って時間領域における関数データに対してフーリエ変換をおこなうことによって周波数領域における関数データを作成する．

Step 5. Step 4 までに得られた時間領域における関数データに対して直流成分が 0 となるようにオフセットを調節する

Step 6. オフセット調整された関数データに対してフーリエ変換をおこなう

Step 7. 実数成分と虚数成分 (または振幅成分と位相成分) の 2 組の関数データを得る

この様にして抽出した 2 組の周波数領域における関数データの周波数軸の間隔は各曲間で同一であるためレジストレーションをおこなう必要がないという利点がある．ただし Step 4 において直線補完をおこなった場合は高周波数領域に雑音が入り込むため，高周波数領域を除いた低～中周波数領域のデータのみを使用する必要がある．

ここで実際に楽曲から関数データを作成した例を示す．例として用いた楽曲は民謡のずいずいずっころばしであり，図 1 の五線譜は主旋律の最初の 2 小節分を表している．また図 2 は主旋律の最初の 4 小節から作成した時間領域での関数データであり，縦軸は MIDI のノート番号，横軸は時間を表している．また補間方法としては直線補間を用いた．

図 3，及び図 4 は周波数領域における 2 組の関数データであり，図 3 が実数成分，図 4 が虚数成分の関数データである．なお，図より振幅の小さい $w \geq 15$ あたりの帯域にあるデータは線形補間により生じた雑音と考えられるため，特徴抽出をおこなう場合はそれよりも低い帯域のデータを用いる必要がある．

4 適用例

ここで適用例としていくつかの楽曲から作成した関数データを元にして判別分析をおこなった例を示す．



図 1: 五線譜上の主旋律

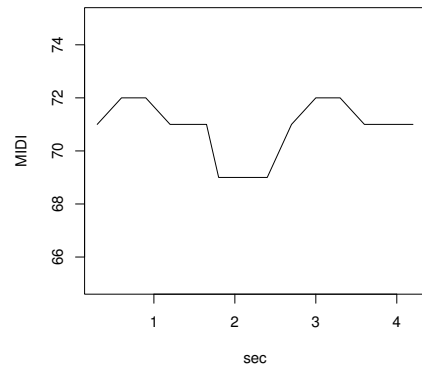


図 2: 主旋律から作成した時間領域の関数データ

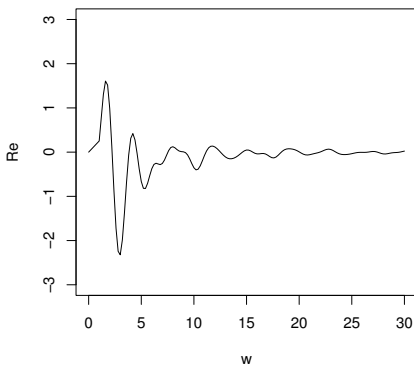


図 3: 周波数領域の関数データ (実数成分)

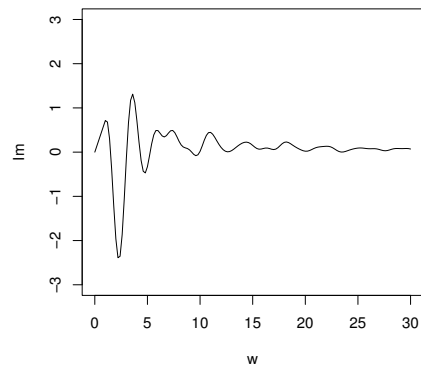


図 4: 周波数領域の関数データ (虚数成分)

はじめに被験者が各楽曲のイントロ部分を聞いてその曲が好きか嫌いかの嗜好性アンケートをおこなった。次に各楽曲のイントロ部分から関数データを作成して各周波数帯域ごとに好きの群と嫌いの群についての判別分析をおこなった。なお今回の実験では 14 曲の楽曲を使用した。この判別分析の結果を図 5 に示す。図 5 の横軸は周波数、縦軸は各周波数における判別分析の正答率である。

図 5 より、正答率の高い周波数帯域には被験者の嗜好性 (好き・嫌い) をはっきりと区別する特徴成分が含まれていると思われる。そこで $w = 5$ において判別分析をおこなった結果を図 6 に示す。図 6 の横軸は実数成分、縦軸は虚数成分であり、○ は好き、× は嫌いを表す。群平均は好きの群が $(-0.4954459, 1.25989129)$ 、嫌いの群が $(0.5684148, 0.06770057)$ である。この結果によれば群平均の振幅は嫌いの群がやや小さくて位相はほぼ反転していることが分かる。従ってある楽曲の $w = 5$ 周辺の帯域における関数データの、特に位相に関して被験者の嗜好性が特徴として現れていると考えられる。

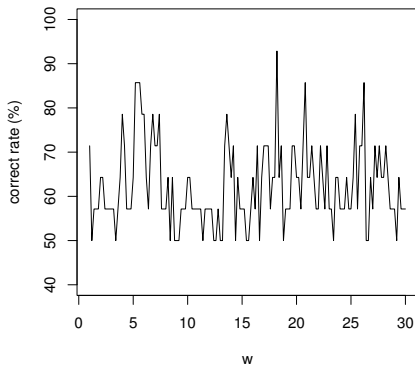


図 5: 正答率

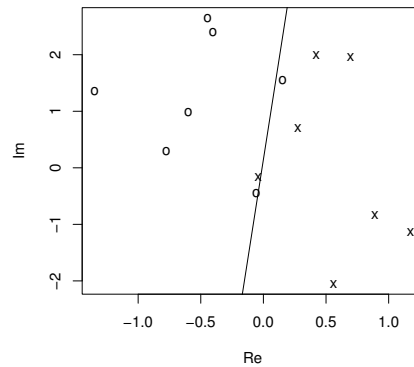


図 6: $w = 5$ における判別分析結果 (横軸が実数成分, 縦軸が虚数成分)

5 まとめ

本稿では楽曲の旋律を時間領域における関数データとみなし, さらにフーリエ変換をおこなうことによって 2 組の周波数領域における関数データを作成してそれらを楽曲の特徴量として使用することを提案した. また適用例として関数データに対して判別分析をおこなって被験者の嗜好性 (好き・嫌い) 情報の特徴抽出をおこなった.

なお今回用いた 2 組の関数データだけでは複雑な特徴を抽出するには明らかに不足しているため, 実際には音符の構成情報や局所パターン情報などの局所的, 大域的なデータも含めて総合的に特徴を抽出することが望ましいと思われる. また今回は適用例として嗜好性の様に主観的な特徴を使用したため, 客観的な特徴も関数データによって抽出可能であるかを評価することが今後の課題となる.

参考文献

- [1] 帆足啓一郎, 上月勝博, 菅谷史昭 (2005), 楽曲配信サービスを支える音楽情報検索技術, 電子情報通信学会誌, Vol.88, No.7, pp.529-534.
- [2] 辻康博, 星守, 大森匡 (1997), 曲の局所パターン特徴量を用いた類似曲検索・感性語による検索, 信学技報, SP96-124, pp.17-24.
- [3] 田村理遊, 池田剛, 小谷義行, 但馬 康宏 (2006), 隠れマルコフモデルを用いた副旋律推定, 情報処理学会研究報告, 2006-MUS-64, pp. 17-22.
- [4] Ramsay.J.O. and Silverman.B.W. (1997), *Functional Data Analysis*, Springer.
- [5] 下川真由子, 水田正弘, 佐藤義治 (2000), 関数データ解析における関数回帰分析の拡張, 応用統計学, Vol.29, No1, pp. 27-39.