複数のパートに分散したメロディを抽出するための一手法*

芹澤裕子　鈴木伸崇　佐藤洋一郎　早瀬道芳†

岡山県立大学大学院　情報系工学研究科‡

1. はじめに

標準MIDIファイルは，通常，複数のパートから構成され，メロディと共に伴奏が記録されている．このようなデータに対してメロディ検索を行うためには，伴奏を除外してメロディのみを抽出する必要が生じる．

メロディ抽出の対象となる楽曲が歌謡曲である場合，メロディパートが単一のパートを占有していることが多いため，メロディを担当する単一のパートを選出することにより，曲全体のメロディを抽出することが可能である [1]．しかし，この手法は，クラシック音楽などのようにメロディが複数のパートに分散している楽曲に対しては必ずしも有効でない．そこで，本研究では，メロディが複数のパートに分散している標準MIDIファイルからメロディを抽出する手法を提案する．

2. 本手法の概要

本手法では，まず，楽曲を複数のフレーズに分割する（各パートは同じ位置で区切られる）．次に，各フレーズの各パートに対して評価値を計算し，最も評価値（後述）の高いパートをそのフレーズにおけるメロディとする．以下，フレーズへの分割方法と評価値の計算方法について説明する．

3. フレーズへの分割

以下の手順で楽曲をフレーズに分割する．

1. 各パートに対して，「隣接する2音の音高差と発音間隔の積」を求める．それらの中で最大値をとる点で，全てのパートを区切る．
2. 区切った範囲が四分音符16拍以下となるまで，(1)の処理を再帰的に行う．

フレーズの長さを16拍以下としたのは，楽曲中において音楽的に意味を持つメロディの長さの単位が，16拍となることが多いからである．

4. 評価値の計算

まず，音長の正規化を行う．標準MIDIファイルは人間が楽器を演奏することにより生成されることも多いため，打鍵のタイミングが楽譜とは微妙にずれている場合がある．このため，同じ種類の音符でも実際の音長が異なる場合がある．この音長のずれを修正するため，標準MIDIファイルの解像度より求めた四分音符の長さを基に，各音の音長を，楽譜上で最も近い音符の長さに置き換える．

次に，各フレーズの各パートに対して評価値を計算する．以下の5個の評価項目に関して評価値を加算・減算していく（評価値の初期値は0とする）．

平均音高

音の高い音符を多く含むパートはメロディとして認識される可能性が高いと考えられる．そこで，各フレーズの各パートに対して，次のような平均音高を求め，その値が最も大きいパートの評価値を1加算する．

$$\text{平均音高} = \frac{\text{各音の音高の総和}}{\text{音数}}$$

平均音量

まず，各パート内に和音が存在する場合，最も音量の大きいもの（同音量の場合は，最も音高が高いもの）を選択するという処理（単音化）を行っておく．

次に，平均音量を求める．メロディであるパートは，他のパートと比較して音量が大きい場合が多い．そこで，各フレーズにおいて平均音量が最も大きいパート（平均音量が等しい場合は，平均音高が高いパート）の評価値を1加算する．ここで，フレーズ長を無視して音量の平均を取ると，フレーズ内に休符が多く，わずかに存在する音符の音量が大きい場合，平均音量は大きくなる．しかし，休符が多いパートがメロディパートである可能性は低いので，このようなパートの評価値を高くするのは不適当である．そこで，平均音量を以下のように定義する．

$$\text{平均音量} = \frac{\text{各音の音量と音長の積の総和}}{\text{フレーズ長}}$$

---

*A method for extracting a single melody line from a music containing multiple parts

†Yuko Serizawa, Nobutaka Suzuki, Youichirou Sato, Michiyoshi Hayase

‡Graduate School of Faculty of Computer Science and System Engineering, Okayama Prefectural University

これにより，パート内に休符が多い場合，平均音量が小さくなる．

発音率

フレーズ内で音符が占める割合が少ないパートは，メロディである可能性が低い．このようなパートを識別するため，以下のような発音率を定義する．

$$発音率 = \frac{各音の音長の総和}{フレーズ長}$$

本手法では，発音率の閾値を 0.5 とし，発音率が 0.5 未満のパートの評価値を 3 減算する．

楽器指定

音階を持たない打楽器によるパートは，メロディとなる可能性が極めて低い．したがって，各パートの楽器を調べ，音階を持たない楽器の場合はその評価値を $-\infty$ とする．GM では 128 種類の音色とそれに対応する音色番号が定義されているので，ProgramChange という MIDI メッセージの値を調べることにより，各パートで使用されている楽器を特定できる．

音高パターン

多数のパートが同じ音高で同時に演奏した場合，それらのパートは他のパートよりもメロディとして認識されることが多い．これは，多数のパートが同時に演奏する音高の音は，他の音に比べて音量が大きくなるためである．そこで，音高パターンの類似したパート同士にグループ分けし，パート数が最大であるグループに属するパートの評価値を 1 加算する．グループに分ける手順を以下に示す．

1. フレーズ中の各パートに対して，各音の音高をオクターブ単位の音高差を無視して 12 音に分け，各音高の発生回数を求めて，ヒストグラムを生成する．
2. ヒストグラムを比較することにより，各パート間の音高パターンの類似度を求める．
3. 音高パターンの類似度が 0.95 以上となるパートは同じグループとみなす．

音高パターンの類似度は，2 つのヒストグラムをベクトルとみなし，コサイン相関値を計算することにより求める．また，和音が存在する場合は，単音化してからヒストグラムを生成する．

リズムパターン

多数のパートが同じリズムパターンで演奏した場合，それらの音高パターンが同じであればメロディである可能性は高くなる．また，あるパートのリズムパターンが他のどのパートのそれとも異なる場合，そのパートは目立って聞こえるので，メロディである可能性がある．そこで，各パートを類似したリズムパターンを持つパート同士にグループ分けし，リズムパターンの最大グループと音高パターンの最大グループに共に属するパートの評価値を 1 加算する．また，リズムパターンが他のどのパートとも異なるパートが存在する場合，その評価値を 1 加算する．

リズムパターンは，各音符の発生回数によるヒストグラムとする．音符の種類は，全音符以上，全音符，付点二分音符，二分音符，付点四分音符，四分音符，付点八分音符，八分音符，付点十六分音符，十六分音符，三十二分音符の 11 種類とする．グループ分けには，ヒストグラム間のコサイン相関値を用いる．

5. 実験結果

クラシック音楽 10 曲を対象にして，メロディの抽出実験を行った．5 人の被験者にメロディパートを選択してもらい，最も多くの被験者が選択したパートを正解とみなして本手法の出力と比較した．メロディパートの抽出成功率を四分音符単位で求めたところ，約 82.9%のメロディパートを抽出することに成功した．

メロディが正確に行えなかった部分のうち，比較的多いのが複数のパートの演奏内容が対等な関係にある部分（伴奏とメロディが明確に区別できないような部分）である．このような部分に対しては被験者の意見が複数に分かれることが多く，本手法の出力と正解とは必ずしも一致していなかった．

6. まとめ

本稿では，メロディと伴奏が混在する標準 MIDI ファイルからメロディのみを抽出する手法を提案した．そして，実験の結果，比較的高い割合でメロディを抽出可能であることを示した．本手法を発展させることにより，クラシック音楽などを対象としたメロディ検索に応用できると考えられる．

参考文献

[1] 鷲坂光一，“標準 MIDI ファイルからのメロディの自動抽出法”，情報処理学会技術研究報告「音楽情報科学」，No.005–2，pp.7–12，1994.

[2] 関本陽子，野池賢二，乾伸雄，野瀬隆，小谷義行，西村恕彦，“楽譜情報からの主旋律判定関数の生成”，情報処理学会第 52 回全国大会論文集，pp.435–436，1996.