3L-02

日本語楽曲の旋律と歌詞のアクセントの関係分析のための自動対応付け

藤村　春花†　　岩野　公司†

東京都市大学†

1. はじめに

これまで，楽曲のメロディ（旋律）と歌詞を読み上げた時の声の高低パターン（アクセント）の類似度を分析する研究が行われている（[1]など）．分析で得られた知見は，時代による曲調変化やヒット曲の法則性の調査などに有用と考えられるが，一般的な知見を得るためには大量の楽曲データで分析を行う必要がある．しかし，各音符に歌詞が割り当てられている楽譜データを人手で大量に用意することは難しい．

そこで本研究では，楽曲の MIDI データと歌詞テキストを独立に与えた状況で，旋律とアクセントの類似度を自動的に分析するシステムの構築を目指す．まず，A メロ，サビといったパートごとに分かれたメロディと歌詞に対して，歌詞のどの部分がどの音符に対応付けられるかを自動的に推定する手法の提案を行う．その上で，提案手法を実装した類似度分析システムを構築し，実楽曲の分析を試みる．

2. 旋律とアクセントの類似度分析システム

システムの構成を図 1 に示す．入力の楽曲データはパート単位の「MIDI データ」と主旋律の「トラック番号」，および「歌詞テキスト」である．まず，与えられた MIDI データとトラック番号から，「メロディ情報抽出」により，主旋律の音符や休符の情報を取り出す．歌詞テキストは読点と改行で句ごとに分かち書きされた文章であり，「形態素解析」によって，仮名で表記されたモーラ系列に変換される．さらに「アクセント推定」によって，その歌詞を読み上げた場合のアクセントが推定され，各モーラに対してピッチの高低が付与される．今回，形態素解析とアクセント推定には，OJAD (Online Japanese Accent Dictionary) の韻律予測機能である「韻律読み上げチュータ・スズキクン[2]」の出力結果を利用する．

次に，音符系列と歌詞のモーラ系列を入力として，歌詞の各モーラがどの音符に対応するかを「モーラー音符対応推定」によって求める．

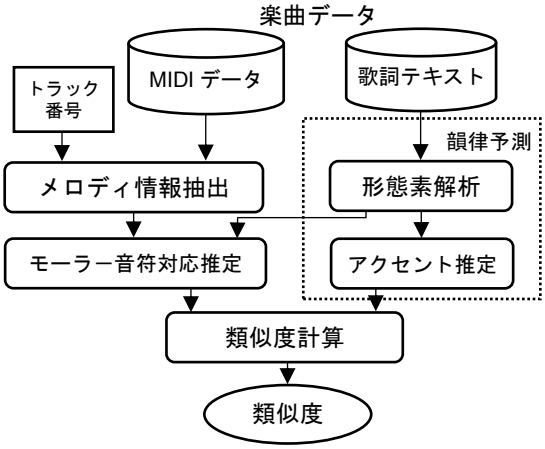

図 1　類似度分析システムの構成

最後に，モーラー音符対応推定結果とアクセント推定結果をモーラごとに比較することで，両者の「類似度計算」を行う．今回は，読み上げ時に「高→低」「低→高」となるピッチの遷移箇所に対し，実際のメロディではどの程度の割合で同じ音程変化となっているかを算出した．

3. モーラー音符対応推定手法

DP マッチングを用いた「モーラー音符対応推定手法」を提案する．横軸にNモーラの歌詞，縦軸にM音のメロディをとった二次元平面を考える．i を歌詞のモーラ番号（$i = 0, \cdots, N-1$），j を音符の番号（$j = 0, \cdots, M-1$）とするとき，各交点(i, j)における，モーラ i と音符 j の適合度合いを表すスコア $s(i, j)$ を以下の式で定義する．

$$s(i,j) = \max \begin{cases} s(i-1, j-1) \times w_1 \\ s(i-1, \quad j \quad) \times w_2 \\ s(\quad i \quad, j-1) \times w_3 \end{cases} \tag{1}$$

w_1，w_2，w_3 は「1 モーラ－1 音」「2 モーラ－1 音」「1 モーラ－2 音」となる対応の生起確率に相当する重みであり，初期値を学習データ中のそれぞれの出現確率 $p_{g_1}, p_{g_2}, p_{g_3}$ に設定する．モーラとして撥音・促音・二重母音が登場した場合には「2 モーラ－1 音」の対応になる可能性が

Automatic alignment for relationship analysis between melody and pitch accent of words in Japanese songs
Haruka Fujimura†, Koji Iwano†, †Tokyo City University

Copyright ©2017 Information Processing Society of Japan.
All Rights Reserved.

大きくなるため，学習データから求めた，それぞれの場合における「2 モーラ－1 音」になる確率 p_b を用いて w_1, w_2 を以下のように変化させる．

$$w_1 = (1 - p_b) \times (p_{g_1} + p_{g_2}) \quad (2)$$
$$w_2 = p_b \times (p_{g_1} + p_{g_2}) \quad (3)$$

また，歌詞の読点や改行の位置では休符が発生しやすいことから，両者の同期性を考慮し，読点と改行をそれぞれ記号化して歌詞のモーラ系列に含め，さらに，一定以上（16 分休符以上）の長さをもつ休符を表す記号も音符系列に含めた上で，以下のルールで重みを設定する．

- 読点・改行と休符が対応する交点：$w_1 = p_{n_1}$, $w_2 = w_3 = 0$
- 読点・改行と音符が対応する交点：$w_2 = p_{n_2}$, $w_1 = w_3 = 0$
- 仮名（モーラ）と休符が対応する交点：$w_3 = p_{n_3}$, $w_1 = w_2 = 0$

以上で定義されるスコアに基づく DP マッチングにより，$s(N-1, M-1)$ を最大とするパスを求め，両者の対応を決定する．

4. 評価実験

4.1 実験データ

評価実験には RWC 研究用音楽データベース[3]から，日本語のみで歌詞が構成されている楽曲を使用する．学習データには，ポピュラー音楽10曲を使用し，評価データには学習データとは別のポピュラー音楽5曲のAメロとサビ，および，童謡5曲のメインパートの計15パートを使用した．p_{g_1}, p_{g_2}, p_{g_3} は学習データを基にそれぞれ 0.968, 0.023, 0.008 と決定し，$p_{n_1}, p_{n_2}, p_{n_3}$ の値は，予備実験の結果から，読点に対して 1.0, 0.9, 0.9，改行に対して 1.0, 0.0, 1.0 と設定した．

4.2 モーラ－音符対応推定手法の評価

提案する対応推定手法の評価結果を表 1 に示す．数値は対応付けられた全ペアのうちの正解の割合（正解率）を示している．正解率は平均91%と高い精度が得られた．童謡で性能が十分に得られていないのは，学習データ（ポピュラー音楽）とスタイルが異なっている（特に，童謡中には「1 モーラ－2 音」の対応が多く含まれている）ことが原因と考えられる．

4.3 類似度の算出結果

提案手法で求められた対応推定結果とアクセント推定の結果から，各モーラに対する音程とアクセントの高低の類似度（一致率）を計算したところ，ポピュラー音楽で52%，童謡で46%となった．人手で求めた正解値はそれぞれ 50%, 48%となっていることから算出結果の誤差は約

表 1　モーラ－音符対応推定性能

ジャンル	正解率
ポピュラー音楽	0.95
童謡	0.64
平均	0.91

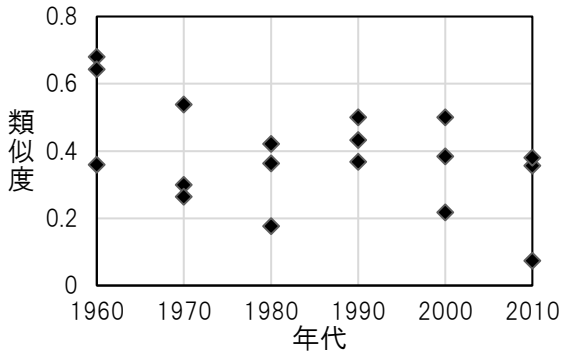

図 2　各年代の楽曲における旋律とアクセントの類似度算出結果

2%で，概ね良好であることがわかった．最後に，本システムによる分析の一例として，1960～2010年代の代表的な歌謡曲を3曲ずつ選び，そのサビ部分の旋律とアクセントの類似度を算出した．図 2 に結果を示す．これを見ると，時代が進むにつれて歌詞のアクセントと一致しないメロディが割り当てられている傾向が読み取れる．

5. まとめ

本研究では，楽曲のメロディと歌詞のアクセントの類似度分析システムの構築を目的とし，モーラ－音符対応推定手法の提案と，システムから得られた類似度の評価を行った．その結果，提案する対応推定手法の精度は約91%，最終的に算出された類似度の誤差は約 2%と良好な性能となることが示された．

今後の課題としては，「1 モーラ－2 音」への対策による対応推定手法の性能改善があげられる．また，構築システムを用いた様々な楽曲に対する類似度分析もさらに進める必要がある．

参考文献

[1] 高木，“歌謡曲における歌詞のアクセントと旋律との関係：「雨」と「霧」を例にして，” 言語文化研究：中部大学女子短期大学紀要３，pp.109-113,1991.

[2] 峯松，“日本語音声・テキストコーパス情報処理に基づくオンライン韻律教育インフラの構築，” 音声研究，vol. 19, no. 1, pp. 18-31, 2015.

[3] 後藤他，“RWC 研究用音楽データベース：ポピュラー音楽データベースと著作権切れ音楽データベース，” 情報処理学会音楽情報科学研究会研究報告，2001-MUS-42-6, vol.2001, no.103, pp. 35-42, 2001.

Copyright ©2017 Information Processing Society of Japan.
All Rights Reserved.