

入力文書の印象と感情に基づく楽曲提供の一手法

菅野 沙也 (指導教員：伊藤 貴之)

1 概要

我々は小説などの文書の印象を様々な形で表現する。特にレビューを書くなどの形で別の文章を用いたり、イラスト化するなどの視覚的手法を用いたりすることが多い。しかし、これらの手段には作品のネタバレをさせたり、登場人物の容姿や場所のイメージに関する先入観を与えすぎたりする可能性がある。そこで文書の印象を表現する別の手段として、昔から我々の生活や感情表現に欠かせない要素である「音楽」で文書の印象を表現することを提案する。音楽にはイメージ誘導効果・行動誘導効果・感情誘導効果があり、これは音楽への興味の有無に関わらず、すべての人に効果があると行動心理学の観点 [1] から語られている。さらに、文学と音楽という異なる種類の芸術をリンクさせることで、新たな感動や刺激が得られることも期待できる。このような効果は、オペラやバレエ音楽などで既に実現されている。しかし楽曲の制作には専門的知識や経験が必要である。特に、ある特定のイメージを基に楽曲を制作するのは困難であり、誰でも気軽に実現できるとは言いがたいのが現状である。

そこで本研究では、文書を入力するとその印象や感情に基づいた楽曲を提供する一手法を提案する。本手法のようなシステムにより、音楽を用いて文章作品の印象や感情の間接的な伝達が促進できると期待できる。

2 関連研究

文書から楽曲を生成する研究として、物語の論旨や文法の解析結果に基づいて楽曲を生成する手法 [2] がある。しかしこの手法は、読者が受けた印象や感情に基づいたものではない。一方で、文章の印象や感情に基づいた楽曲生成手法として、文章の印象値からの機械的な音列生成によって楽曲を自動生成する手法 [3] がある。しかしこの手法では、機械的な音列生成であるがゆえに、音楽的表現に限りがあり、また例えば自分の好みの音楽的素材を採用することが難しい。それに対して本研究では、あらかじめ用意したコード進行とリズム素材のマッシュアップによって楽曲を生成する、という点が関連研究と異なる。

3 提案手法

本手法は大きく分けて3つの処理段階で構成される。具体的には、

事前調査：感性語と音楽的特徴量の選出

学習：コードとリズムの印象の関係性算出

対話処理：ユーザーの文書入力に対する楽曲生成の3段階である。詳細について以下に論述する。

3.1 事前調査

まず、文書解析の基準となる感性語対と、コードやリズムの選出基準となる音楽的特徴量を選出する。我々は自身の主観に基づいて、図1に列挙する感性語対と音楽的特徴量を候補にあげた。

コード(和音)進行

感性語候補

- ・明るいー暗い
- ・軽いー重い
- ・楽しいー切ない
- ・派手ー地味
- ・軽快ーしっとり
- ・元気なー落ち着いた

音楽的特徴量候補

- ・音程の平均的な高さ
- ・音の平均的な広がり
- ・音の平均的な厚み
- ・不協和音度
- ・メジャー/マイナー/セブンス/メジャーセブンス/マイナーセブンスの割合

リズム進行

感性語候補

- ・速いー遅い
- ・軽いー重い
- ・静かーうるさい
- ・派手ー地味
- ・穏やかー激しい

音楽的特徴量候補

- ・ドラムの各部位を叩いた割合
- ・全音符数
- ・16分音符/8分音符/4分音符以上の割合
- ・3連符の割合

図1: 感性語対と音楽的特徴量の候補

コード(和音)進行

決定した感性語

- ・明るいー暗い
- ・楽しいー切ない
- ・軽快ーしっとり
- ・元気なー落ち着いた

決定した音楽的特徴量

- ・音程の平均的な高さ
- ・音の平均的な広がり
- ・音の平均的な厚み
- ・不協和音度
- ・メジャー/メジャー7th/マイナー7thの割合

リズム進行

決定した感性語

- ・速いー遅い
- ・軽いー重い
- ・派手ー地味

決定した音楽的特徴量

- ・タム/スネア/金物/バスドラムの割合
- ・全音符数
- ・16分音符の割合
- ・3連符の割合

図2: 決定した感性語対と音楽的特徴量

続いて、感性語対と音楽的特徴量の選出のためにユーザーアンケートを実施した。コードとリズムの音源を聴き、それぞれ候補となる感性語対の度合いを5段階評価で回答してもらった。この回答をもとに、感性語対に関する音源Nの適合度と音源Nについての音楽的特徴量の相関性を相関係数により算出し、どの感性語対とも相関性が小さい音楽的特徴量、またはどの音楽的特徴量とも相関性が小さい感性語対を対象から除外した。以上の結果として、コード進行およびリズム進行の感性語対と音楽的特徴量を図2のように選出した。

続いて、日本語の各内容語(名詞、動詞、形容詞、副詞)が、選出した各感性語対のいずれの感性に近いかを数値化する。具体的には感性語対の一方を+1に、もう一方を-1に対応させ、-1から+1の実数値を各内容語に対し算出する。この算出には感情極性抽出手法 [4] を適用し、各感性語対を種単語として用いた。

3.2 学習

続いて、コードおよびリズムの音楽的特徴量とそれに対する各ユーザーの印象の関係を学習する。現時点での我々の実装では、各ユーザーにサンプルコードおよびサンプルリズムを聴いてもらい、3.1節で選出した感性語への適合度を回答してもらい、以後、この適合度を印象値と称する。また、印象値を推測するための式を音楽的特徴量の線形結合(図3参照)と仮定し、重回帰分析を用いてこの式の各係数を算出する。この式を用いてコード進行とリズム進行の印象値を算出し、

コード進行
 印象値 $X = a_1 \times [\text{majの割合}] + a_2 \times [\text{M7の割合}] + a_3 \times [\text{m7の割合}] + a_4 \times [\text{音高の平均}] + a_5 \times [\text{音幅の平均}] + a_6 \times [\text{音厚の平均}] + a_7 \times [\text{音厚の平均}]$

リズム進行
 印象値 $Y = b_1 \times [\text{全音符数}] + b_2 \times [\text{16分音符の割合}] + b_3 \times [\text{3連符の割合}] + b_4 \times [\text{金物の割合}] + b_5 \times [\text{バスドラの割合}] + b_6 \times [\text{タムの割合}] + b_7 \times [\text{スネアの割合}]$

図 3: コード進行およびリズム進行の印象推測のための式

あらかじめ格納する。以上の学習により、ユーザーごとの印象の違いを考慮した楽曲生成が可能になる。

3.3 対話処理

最後に、実際にユーザーが文書を入力すると楽曲を提供する対話処理について論じる。対話処理はさらに文書解析、コードとリズムの選出、楽曲生成の3段階に分けられる。

3.3.1 文書解析

まずユーザーが入力した文書に関して文書解析を適用する。現時点の我々の実装では、文書中でコード・リズムを切り替える点をユーザー自身によってタグ等で指定したテキストファイルを前提とし、これを入力することで場面ごとに文書の印象値を算出する。

我々の実装では、オープンソース形態素解析エンジン MeCab [5] を用いて入力文書を形態素解析し、そのうち名詞、動詞、形容詞、副詞のみを抽出して、感性極性辞書を用いて数値化する。これの分散をとることによって、各感性語対に対する印象値を算出する。

3.3.2 コードとリズムの選出

次に楽曲の素材となるコードとリズムを選出する。ここでは3.3.1項で算出した印象値を多次元ベクトルとして扱い、これと3.2節で算出したコードおよびリズムの印象値を比較して、ユークリッド空間上で最も距離の近いコードおよびリズムを、入力文書の印象に沿った楽曲の素材とする。現時点での実験環境では、コードについては「明るい」「暗い」「楽しい」「切ない」「軽快」「しっとり」「元気な」「落ち着いた」の4次元を採用している。リズムについては「速い」「遅い」「軽い」「重い」「派手」「地味」の3次元を採用している。

3.3.3 楽曲生成

選出したコードとリズムを合成することで楽曲を生成する。我々の実装ではコードとリズムのMIDIファイルはそれぞれ単一のトラックで構成されることを前提とし、この各々のトラックを新しく生成するMIDIファイルの1番目、2番目のトラックにコピーする。現時点での我々の実装ではメロディーを扱っていないが、今後はメロディーの選出とコピーを実装したい。

4 実行結果と考察

本手法によってユーザーAとユーザーBの感性をもとに童話「シンデレラ」を楽曲化した結果を図4に示す。

ユーザーAとユーザーBでは異なる楽曲素材が選ばれており、学習段階の影響により各ユーザーごとの

	ユーザーA	ユーザーB
場面1	chord10.mid, rhythm2.mid	chord8.mid, rhythm12.mid
場面2	chord13.mid, rhythm2.mid	chord5.mid, rhythm15.mid
場面3	chord14.mid, rhythm2.mid	chord5.mid, rhythm12.mid
場面4	chord13.mid, rhythm2.mid	chord5.mid, rhythm12.mid

図 4: 童話「シンデレラ」を楽曲化した結果

印象の違いを考慮した楽曲が生成されていると考えられる。しかし両ユーザーともに異なる場面でありながらも全く同じ楽曲素材が選ばれていることがあり、文書の細かな印象や感情の違いを表現できるような改良が必要である。

5 まとめと今後の課題

本研究では、文書の印象値と音楽的特徴量の比較に基づいて入力文書の印象や感情に基づいた楽曲を提供することで、文書の印象表現を促す手法を提案した。今後の課題は以下の通りである。

- 印象値を推測するための式を再検討する。
- 文書中の場面の区切りの自動認識手法を実装する。
- 文書の展開と楽曲の展開に関連を持たせる。
- 楽曲にメロディーを付与する。
- 学習段階におけるコードおよびリズムの評価方法を再検討する。

6 謝辞

本研究を進めるにあたり、東京工業大学 高村先生には感性極性辞書作成プログラムを提供していただきました。ここに感謝致します。

参考文献

- [1] 斉藤寛, 心を動かす音の心理学, ヤマハミュージックメディア, 2011.
- [2] 遠藤順, 北館拓真, 小方孝, 物語生成システムにおける音楽の生成・表現機構, 日本認知科学会第29回大会, pp. 3-29, 2012.
- [3] 北原和也, 渡邊英徳, 安藤大地, ウェブアクティビティを反映させた楽曲の自動生成手法, 先端芸術音楽創作学会会報, Vol. 1, No. 1, pp. 8-11, 2012.
- [4] 高村大也, 乾孝司, 奥村学, スピンモデルによる単語の感情極性抽出, 情報処理学会論文誌, Vol. 47, No. 2, pp. 627-637, 2006.
- [5] 京都大学情報学研究科 日本電信電話株式会社コミュニケーション科学基礎研究所 共同研究ユニットプロジェクト, MeCab, <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>