

音響特徴と歌詞特徴を用いた楽曲の感情分類手法の提案

今村 環紀[†] 倪 宝栄^{††}

[†] 福岡工業大学大学院 工学研究科電子情報工学専攻

^{††} 福岡工業大学 工学部電子情報工学科

1. はじめに

近年、音楽ストリーミングサービス (Spotify、Apple Music など) の普及によって、インターネットを介してユーザーは数千万曲もの膨大な楽曲に対して容易にアクセスができるようになった。ここで、この膨大な楽曲の中からその時々においてユーザーに最適な楽曲を推薦する、楽曲推薦システムが注目されている。これは主に楽曲の音響特徴やジャンルなどのメタデータの類似性により推薦を行う[1]が、楽曲の持つ感情をメタデータとして扱うことができれば、ユーザーの気分によって聴きたいムードを持つ楽曲を推薦する推薦システムへと利用することができる。楽曲の感情分類、推薦に主に利用される特徴として、歌詞と音響特徴の2点が挙げられる。同分野の研究では主に歌詞のみ、音響特徴のみで行われているものが多く、本研究では楽曲推薦システムの精度向上のために、この2種類の特徴を基にして楽曲にムードタグをつける事を目標とする。

2. 感情モデル

感情モデルは これまでに多種多様なモデルが提案されている。過去に楽曲推薦の分野でしばしば用いられている事[1]から今回ラッセルの円環モデル[2]を参考にしている。縦軸覚醒の度合いを表す Arousal、横軸快-不快を表す Valence の2軸による二次平面(AV 平面)ですべての感情を表す事ができる。

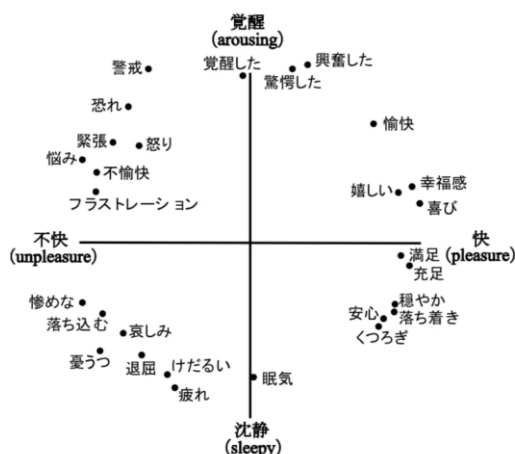


図1. ラッセルの円環モデルの略図

3. 音響特徴による分類

音響特徴量によるムード分類は、今回主に Spotify の API から取得できる特徴量を元にする。Spotify から取得できる特徴量のうち、先程の感情モデルの Arousal, Valence に対応する特徴量として、Energy と Valence がある。オーディオ信号から得られる7つの特徴量 (acousticness, danceability, instrumentalness, liveness, loudness, speechiness, tempo) を用いてランダムフォレストによる各特徴量の寄与率を算出する事で、音響特徴量から Arousal, Valence の値を算出する。

4. 歌詞による分類

今回歌詞による分類は日本語の歌詞の楽曲を対象にする。英語の歌詞でムード分類を行う場合には、感情ラベルのついたデータセットを教師データとして学習を行うことが可能だが、日本語の場合大規模なデータセットがなく、困難である。よってムード分類の手法については、今回ラッセルの円環モデル内にある、感情単語 (嬉しい、哀しみ等) と歌詞テキストとの類似度計算によって行う。

5. まとめ

各分類の詳しい結果については発表の当日に話す。今回は主に2種類の特徴それぞれによるムード分類に焦点を当てて評価する。今後はこれらの結果を元に音響の特徴と歌詞の特徴の2種類によるマルチモーダルなムード分類の手法について考えていく。

参考文献

- [1] 奥健太, "楽曲推薦システム ープレイリスト, コンテキスト, インタラクション ー" 人工知能誌, vol.34, no.3, pp.300-308, May, 2019.
- [2] Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178