

BGMにおけるタグ情報および音楽的特徴量の類似性の 相関に関する考察

横尾 将人[†] 田中 敦[‡]

山形大学大学院理工学研究科^{†‡}

1. はじめに

近年音楽配信サービスの普及が目覚ましい。サービスの利用により、大量の音楽と出会うことができるようになった。しかし、大量の楽曲の扱いがユーザに委ねられているため、自身の好む楽曲を適切に見つけ出すことが難しい。そのため、音楽推薦システムの研究が盛んになされている。中でも、ユーザの取り巻く状況に合致する楽曲の推薦が注目を集めている。小田川らは、ドライブ中において、ユーザが聞きたい楽曲の印象語を入力することで、ユーザの感性に合った楽曲を推薦する技術を提案している[1]。

音楽推薦システムにおいては、コンテンツベースフィルタリングがよく用いられる。音楽内容として、音響信号に基づいたデータとメタデータに大きく分かれる。前者は音楽ファイルを解析することで音楽的特徴量として得ることができる。後者はアノテーションや、不特定多数のユーザが楽曲に付与したソーシャルタグを参照することで利用できる。

音楽推薦システムにおいて音楽内容が使われるが、推薦理由の明確な提示を行うことができていない現状がある。その多くは、音響信号に基づいたデータからでは推薦結果の分析が困難であるためである。メタデータに基づいた推薦は推薦結果の分析に有用であるが、必ずしも付与されているわけではない。

本研究では、音楽的特徴量とメタ情報との対応の発見により問題を解決できると考えた。その前段階として、音楽的特徴量とメタ情報の相関について考察を行う。今回、ユーザの取り巻く状況に適した楽曲推薦を見据え、メタ情報のなかでも印象語に焦点を当てる。

2. 使用データ

本研究で使用するデータは、フリーBGMダウンロードサイト DOVA-SYNDROME[2]より取得する。DOVA-SYNDROMEは2008年3月よりサービスを開始、著作権フリーなBGMをダウンロードできるサイトである。このサイトでは、楽曲ごとにソーシャルタグが付与されている。

本研究では、2014年8月から2015年9月の一年間を対象として、音楽ファイル(1163曲)とそれに付与された印象語(表1)を収集した。音楽ファイルから、数値解析ソフトMATLABのパッケージの一つであるMIRToolbox[3]を使用し、七つの音楽的特徴量(表2)を算出した。

3. 印象語の音楽的特徴量による評価

各楽曲には印象語がいくつか付与されている。それをもとに、楽曲をクラスタリングする。クラスタリング手法は、教師あり学習の非線形サポートベクターマシン(以下非線形SVMと記述)を用いる。

SVMには教師信号が必要である。教師信号には印象語を基にしたものを使用する。各楽

表1 雰囲気・感情・世界観に属するタグ群

明るい	楽しい	温かい	穏やか	優しい	爽やか	お洒落
力強い	弱々しい	懐かしい	不思議	騒々しい	可愛い	渋い
シリアス	緊張感	懐かしい	寂しい	切ない	冷たい	暗い
悲しい	怒り	恐ろしい	激しい	サイバー	ファンタジー	コメディ
悪意	怪しい	日常	情熱	冷静	淡泊	希望
絶望	虚無					

表2 算出した音楽的特徴量とそれらの説明

分類	音楽的特徴量	説明
Dynamics	RMS	音圧
	Low energy	弱音の割合
Rhythm	Pulse clarity	リズムの明瞭さ
Timbre	Brightness	1.5kHz以上の音の割合
	Roughness	不協和音の度合
	Spectral irregularity	音質の変化の大きさ
Tonality	Mode	MajorとMinorの音量の差

A study on the correlation of the similarity of the tag information and the musical feature value in BGM

[†]MASATO Yokoo, Yamagata University graduate school of science and engineering

[‡]ATSUSHI Tanaka, Yamagata University graduate school of science and engineering

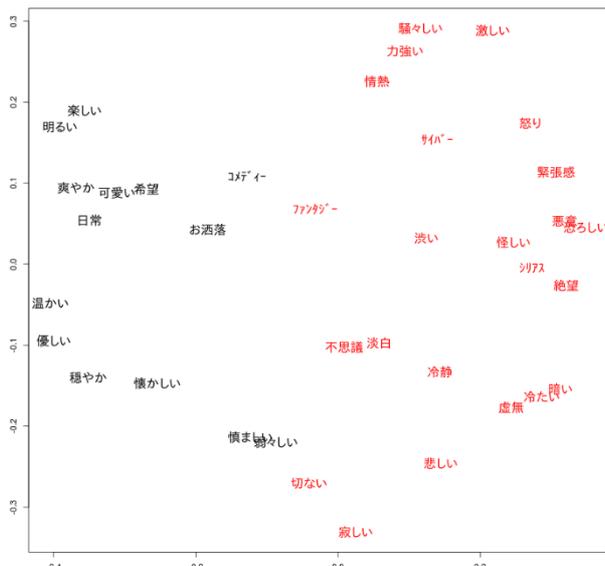


図1 タグのクラスタリング結果

表3 クラスタごとの音楽的特徴用の傾向の推測

Mode	Pulse irregularity
クラスタ1 正の値を取る	比較的低い値を取る
クラスタ2 負の値を取る	比較的高い値を取る

曲には複数の印象語が付与されているため、そのまま教師信号として扱うのは難しい。そこで、印象語をクラスタリングし、各楽曲に新たにクラスを与える。

印象語のクラスタリング流れは次の通りである。

- 印象語間の類似度を Jaccard 係数で定義
- 類似度行列から、多次元尺度構成法により座標を生成
- k-means 法により $k = 2$ としてクラスタリング

印象語のクラスタリングの結果(図 1)、比較的正的意味の印象語群(図 1 の黒字の印象語=クラスタ 1)と比較的ネガティブな意味の印象語(図 1 の赤字の印象語=クラスタ 2)とで分かれた。

各クラスタにおいて、音楽的特徴量の傾向の推測を表 3 に示す。一般的に、Major コードが多いと明るい雰囲気、Minor コードが多いと暗い雰囲気を感じる人が多いとされている。Mode はそれを示す特徴量である。我々はポジティブ・ネガティブという感性と密接に関わっていると考え、表 3 の推測をした。また、Pulse irregularity は楽曲のフレーム間の音の構成の変化の度合を示す特徴量、つまり、周波数スペクトルの分布の変化の度合を示す特徴量である。我々は、ポ

ジティブな感性を楽曲で形作る際には、音の構成を大きくは崩さず、ネガティブはその反対であると考え、推測に至った。

今後図 1 結果をもとに非線形 SVM により、楽曲のクラスタリングを行う予定である。

4. さいごに

DOVA-SYNDROME より一年分のデータを収集し、非線形 SVM の入力に必要なデータを揃えた。印象語のクラスタリングの結果、ポジティブな印象語とネガティブな印象語とで大きく分かれた。今後このクラスタを用いて非線形 SVM により、クラスタリングを行い、音楽的特徴量の傾向を観察する予定である。

印象語のクラスタリングにあたって、今回は二つのクラスタを得ることとしたが、適切なクラスタ数であるとは言えない。したがって、二つのクラスタにおいて、音楽的特徴量の傾向を観察したのち、クラスタ数を変えて同様のことを行う予定である。

参考文献

- [1] 小田川智, 児玉泰輝, 菫山真一, 鈴木康悟, 松下文雄, 塩田岳彦. "楽曲レコメンドシステム". 電子情報通信学会技術研究報告, 応用音響, 2006, Vol.106, No.160, pp.49-53
- [2] "フリーBGM (音楽素材) 無料ダウンロード | DOVA-SYNDROME". DOVA-SYNDROME. <http://dova-s.jp>, (参照 2015-11-25)
- [3] Olivier, Lartillot. MIRtoolbox 1.6.1 User's Manual, 2014-12-7. <https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox/MIRtoolbox1.6.1guide>, (参照 2015-11-25)