

メタ情報と特徴量の関係把握のための楽曲群の可視化

理学専攻 情報科学コース 2040660 渡辺 みどり (指導教員：伊藤 貴之)

1 はじめに

音楽を聴く手段として定額のストリーミングサービスを活用する人が近年増加している¹。音楽ストリーミングサービスにおいて膨大な楽曲の中から任意の楽曲を選択する手段として、ユーザが能動的に好みの楽曲を検索して選曲することや、ユーザの視聴履歴などをもとにシステム側から推薦した楽曲を選択することがあげられる。また、ユーザ自身が楽曲を検索する際に、ジャンルをはじめとした楽曲の分類から探す人も多いと考えられる。このような音楽ストリーミングサービスの普及および楽曲の選択手段の多様化にともない、音楽の自動分類や推薦などの研究が活性化している。

音楽を分類するための指標には、アーティスト名や年代といったメタ情報や、テンポや音量のような音響特徴量など、多様なものが考えられる。これらの指標のうちどれを強く採用するかによって自動分類の結果は変化するが、その最適化は非常に複雑な問題である。音楽分類等における適切な特徴量や次元削減手法の検討や新しいアプローチは数多く報告されている [1][2] が、それらを包括的に調査する研究は少ない。包括的な調査の一手段として、楽曲群の特徴量やメタ情報を同時に可視化することが有効であると考えられる。

そこで本研究では、楽曲分類に効果的なメタ情報、音響特徴量、機械学習手法、可視化手法を包括的に検討するために楽曲群を可視化する。具体的には、あらかじめメタ情報のタグをつけた楽曲群から音響特徴量を算出し、得られた音響特徴量に対して次元削減を適用して可視化する。そして可視化結果を観察することで、メタ情報と音響特徴量の新たな関係を見出すことを目指す。このような可視化を実現することで、ヒット曲の傾向などを探索的に分析して能動的に理解することが容易になる。また、それを活用して音楽制作者・販売者がよりターゲットに合わせた楽曲を生み出すことを支援できると考えられる。

本論文では、テンポや音量平均などの音響特徴量、機械学習による音響特徴量を算出し、年代やアーティストのメタ情報と合わせて可視化した結果を示す。本章以降で本研究が想定する処理の詳細を示し、可視化結果の例を示して考察を述べる。

2 提案手法

図1に本手法の流れを示す。本研究では、年代やアーティストなどのメタ情報がタグとして付与されている楽曲群の音響データを入力情報とする。それらの楽曲について、ツールや機械学習を用いて音響特徴量を算出する。そして得られた特徴量に次元削減を適用し、

メタ情報と合わせて散布図によって楽曲の分布を可視化する。

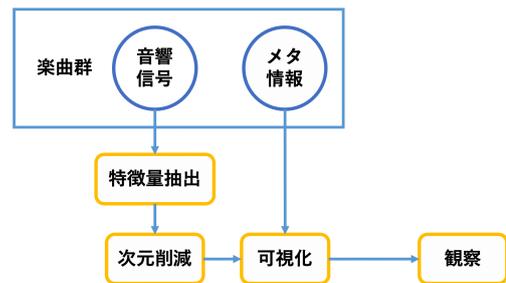


図 1: 提案手法の概要

2.1 メタ情報

本論文では、各楽曲に対してタグとして付随させる情報をメタ情報と呼ぶ。楽曲に対するメタ情報の例として、アーティスト、曲名、作曲家、発売日、歌詞などがあげられる。本論文では年代、アーティストをメタ情報として可視化した結果を紹介する。

2.2 特徴量抽出

音響信号として録音された楽曲データから、ツールや機械学習を用いて数値化された音響特徴量を算出する。RP_extract²を用いて算出した各楽曲の特徴量を利用する。RP_extract は、リズムパターン (RP)、リズムヒストグラム (RH)、統計スペクトル記述子 (SSD) の3種類の特徴量を算出する。本論文では、各クリティカルバンドのビートやエネルギーの発生を統計的に測定して記述する SSD を次元削減し可視化した結果を紹介する。

2.3 可視化

2.2節で算出した特徴量に次元削減を適用して2次元空間に写像し、散布図として楽曲の分布を表示する。メタ情報を基準として散布図上の点群を色分けすることで、メタ情報と特徴量の関係を把握できるようにする。本論文では、t-SNE で次元削減した結果を紹介する。

2.4 観察

2.3節で論じた可視化の結果を観察する。特に以下の3点を重視する。

- ・同一のタグを有する楽曲が同じクラスタに分類されるか。
- ・どの特徴量や機械学習を用いると明確な傾向を持つ

¹<https://www.riaj.or.jp/f/pdf/report/mediauser/softuser2019.pdf>

²<http://ifs.tuwien.ac.at/mir/musicbricks/#RPextract>

た分類を実現できるか。

・分類が困難な楽曲にはどのようなものがあるか。

3 実行例と考察

1986年から2018年までの日本のヒット曲1315曲からなる楽曲群に提案手法を実行した。RP_extractで算出した特徴量SSDをt-SNEで次元削減し、散布図として表示した。年代およびアーティストによる可視化の凡例をそれぞれ表1,2に、可視化結果の散布図を、それぞれ図2,3に示す。

アーティストによる色分けでは、楽曲群のうち、18曲以上が存在する10のアーティストについて色をつけ、残りをその他としてまとめて可視化した。年代で色分けした可視化結果は、年代の古い楽曲(青)から新しい楽曲(赤)へのグラデーションに近い構成をしていることがわかる。アーティストで色分けした可視化結果では、Mr.Children(黄緑)に特に明確なクラスタが見られる。その他、ZARD(茶)、AKB48(青)、EXILE(オレンジ)などの一部の楽曲もクラスタを形成していることが見受けられる。これによって、楽曲の年代間およびアーティスト間では、ビートやエネルギーにおいて特徴の違いがあることがわかった。

さらに、具体的にどのような違いがあるかを確認するため、Librosa³を用いて特徴量を算出し、個別に可視化した。その結果、年代ではrms energy(音量の二乗平均平方根値)に、アーティストではspectral flatness(周波数別の音量の均一さ)に特徴があることを発見した。各楽曲のrms energyの平均値は、概ね年代が上がるほど大きくなった。音量の最大値は過去の楽曲も最近の楽曲もほぼ同一であるため、過去の楽曲ほど音量の変化に富み、最近の楽曲ほど音量が均一化されていることを示す。この原因として、使用される楽器や機器の変化、屋外で聴くことを意識したマスタリング工程などが考えられる。各楽曲のspectral flatnessの分散値は、Mr.Childrenや浜崎あゆみなど一部のアーティストの値が低くなっていた。これらのアーティストの楽曲は周波数別の音量の均一さが曲を通して低かった。そのため、楽器編成や録音手法に一定の特徴があることが示唆される。

■	1986年-1991年	■	2002年-2007年
■	1992年-1996年	■	2008年-2012年
■	1997年-2001年	■	2013年-2018年

表1: 年代による色分けの凡例

■	AKB48	■	工藤静香
■	安室奈美恵	■	ゆず
■	Mr.Children	■	BUMP OF CHICKEN
■	SMAP	■	ZARD
■	浜崎あゆみ	■	その他
■	EXILE		

表2: アーティストによる色分けの凡例

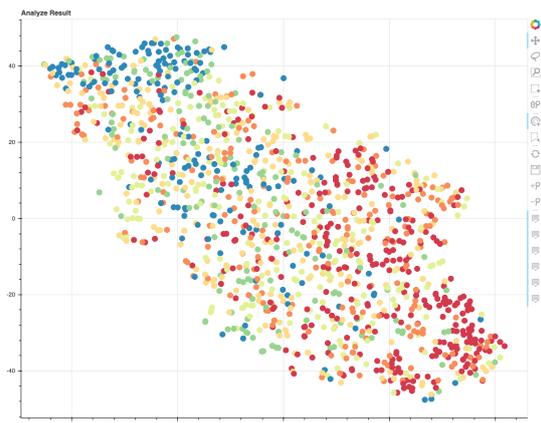


図2: 年代で色分けした結果(凡例は表1)

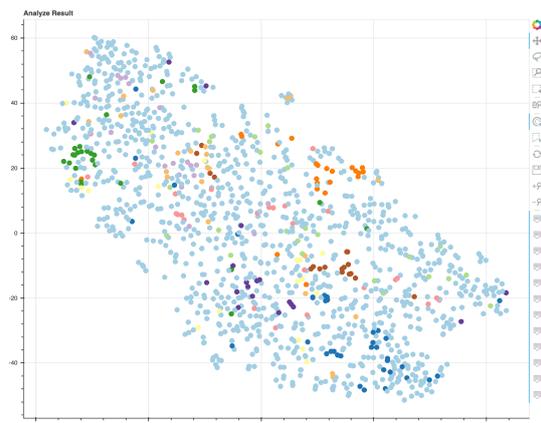


図3: アーティストで色分けした結果(凡例は表2)

4 まとめ

本研究では、メタ情報を付与した楽曲群について、機械学習を用いて特徴量を算出し、メタ情報と特徴量の分布を可視化してその結果を観察した。そして、1980年代から2010年代の各時代の楽曲と、特定のアーティストの楽曲には、ビートやエネルギーに特徴があることを発見した。さらに、これらの特徴の中でも特に、年代間ではrms energyに、アーティスト間ではspectral flatnessに特徴的な差があることを示した。

謝辞: 本研究にあたり、楽曲群データを提供していただいた株式会社レコチョク様、ご協力いただきましたSepteni Japan株式会社の大矢隼士様に感謝申し上げます。

参考文献

- [1] Matteo Lionello, Luca Pietrogrande, Hendrik Purwins and Mohamed Abou-Zleikha, "Interactive Exploration of Musical Space with Parametric t-SNE," 15th Sound and Music Computing Conference, 2018.
- [2] Kin Wai Cheuk, Yin-Jyun Luo, Balamurali B, T, Gemma Roig and Dorien Herremans, "Regression-based music emotion prediction using triplet neural networks," IJCNN 2020, 2020.

³<https://librosa.org/doc/latest/index.html>