

メタ情報と特徴量の関係把握のための楽曲群の可視化

渡辺 みどり[†] 黒子 なるみ^{†*} 大矢 隼士[‡] 伊藤 貴之[‡]

[†]お茶の水女子大学 〒112-8610 東京都文京区大塚 2-1-1

[‡]Septeni Japan 株式会社 〒160-6128 東京都新宿区西新宿 8-17-1

E-mail: [†] {watanabe.midori, kuroko.narumi, itot}@is.ocha.ac.jp, [‡] hayato.oya@septeni.co.jp

*現在株式会社 NTT ドコモ所属

あらまし 音楽の自動分類や推薦に関する研究やサービスが近年活発化している。音楽の分類や推薦には、どのようなメタ情報や音響特徴量がその精度に強く寄与するか、どのようにその有用性を定義するのか、といった点が不明瞭なことが多く、その解明は重要な課題である。そこで我々は、音楽分類タスクに効果的なメタ情報・音響特徴量・機械学習手法・可視化手法を検討すること、音響特徴量とメタ情報の新たな関係を可視化により発見できるか否かを実証することを目標として、メタ情報を付与した楽曲群の可視化に取り組んでいる。具体的には、音楽解析ツールや機械学習を用いて楽曲群の音響特徴量を算出し、音響特徴量とメタ情報の分布を可視化することでその結果を議論する。本報告では、年代、作曲家、アーティストのメタ情報と、エネルギー平均や周波数別エネルギーなどの音響特徴量、機械学習による音響特徴量を用いた実験結果を紹介する。

キーワード 可視化, 音楽情報処理, 音響特徴量, メタ情報

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。

例えば Spotify[2]では、ユーザの行動データから好みの楽曲を推測する協調フィルタリングに加えて、音響特徴量による楽曲分類や類似性をもとにした推薦を組み合わせて楽曲を推薦している。「楽曲をプレイリストに保存する」、「アーティストページに訪問する」といった行動からそのユーザの好みの楽曲リストを作成し、好みの楽曲リストが似たユーザの行動履歴を参照して楽曲を推薦するという協調フィルタリングは、他のストリーミングサービスでもよく利用される手法である。この手法では、そもそもユーザ自身が楽曲を検索し選曲する過程が極めて重要であるが、検索に利用される情報はアーティストやアルバム、ジャンルのようなメタ情報をもとにした分類である場合が多い。また、行動履歴のみでは新しいアーティストの楽曲が推薦されにくくなってしまう問題を回避するために、Spotifyでは、楽曲の音響特徴量を算出してニューラルネットワークを適用し、楽曲間の類似性も考慮した上で音楽を推薦している。このように、音楽ストリーミングサービスにおける楽曲の推薦や分類にも、数多くのメタ情

報や音響特徴量が使用されている。

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。前述したように、音楽推薦システムにおいて多く活用されている協調フィルタリングでは、ユーザの操作履歴などをもとにする場合が多い。さらに、音楽ストリーミングサービスでのユーザの操作の動機には、音楽分類を利用していることが多いと考えられる。以上のことから本研究では、主に音楽分類タスクにおける可視化を想定する。

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。さらに、自動分類がどのように成功もしくは失敗しているかを観察することも課題となる。これらを定性的に判断するには、楽曲分類結果の可視化が有効であると考えられる。音楽分類を目的とした可視化の研究は既に多数報告されている[3][4]が、それらは特定の音響特徴量や機械学習手法によるものが多く、複数のメタ情報や音響特徴量、機械学習手法を組み合わせる包括的に実験された研究はまだ少ない。また、数ある可視化手法のうちどれが効果的であるかも明確になっていない。

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

は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発見することを目指す．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．

本報告では、テンポやエネルギー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、年代や作曲家、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．

本報告の構成は以下の通りである．2 章では関連研究について述べる．3 章では提案手法について、4 章では提案手法の実行例と考察について述べる．そして、5 章で本報告のまとめと今後の課題について述べる．

2. 関連研究

本章では、音楽分類の研究と、音楽可視化の研究を関連研究として示す．

2.1 音楽分類

Burred ら[5]は、音楽ジャンルの自動分類について階層的アプローチを提案した．音響信号から多くの音響特徴量を算出し、ノイズや帯域幅の影響を受けやすい音響特徴量を破棄するなど、音楽分類に効果的な音響特徴量を検討した．提案したアプローチでは直接的なアプローチと同様に音楽を分類できることを示しているが、分類結果を可視化するなどの視覚的表現は示していない．

Cheuk ら[6]は、トリプレットニューラルネットワーク(TNN)を用いた回帰ベースの音楽感情予測手法を提案した．楽曲群の感情データセットに対し主成分分析(Principal Component Analysis, 以下 PCA), オートエンコーダー(AE), TNN を適用した次元削減により、TNN が感情回帰タスクに効果的であることを示した．

Park ら[7]の研究では、メタ情報や音響特徴量から主観的な要素を省き、自動で客観的に音楽を判別している．Park らは、アーティストラベルを用いた教師あり機械学習モデルを提案し、提案手法を応用することでジャンルや性別にもとづいた分類が可能であることを示している．本研究においても、客観的な音響特徴量やメタ情報を使用する．

これらのように、音楽分類等における適切な音響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、楽曲群の音

響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．

2.2 音楽可視化

Lionello ら[8]は、34 次元の音響特徴量を付与した楽曲群にパラメトリック t-SNE(t-distributed Stochastic Neighbor Embedding)を適用し、可視化した．高次元データに対してパラメトリック t-SNE が PCA よりも有用であることを示した．

Flexer[9]は、複数の次元削減手法とハブネス削減方法などを組み合わせ、高次元の音楽類似空間の可視化におけるハブ・アンチハブの悪影響を調査した．共有最近傍法と t-SNE の組み合わせが最もハブネスの影響を小さくできることを示した．

しかし、これらの研究では具体的に用いた音響特徴量に関しては言及されていない．

2.3 Visual Analytics における散布図の活用

可視化技術と分析技術を組み合わせてデータの傾向や特徴を発見する Visual Analytics に関する研究は従来から活発に研究されており、その中で散布図を活用したシステムも多く発表されている．一例として Wilson ら[10]は、高次元データから散布図を作成し、さらにクラスタごとの違いを棒グラフやヒートマップを用いて分析するツールを提案した．クラスタの発見に対する散布図の利用や発見された特徴に対して他の可視化手法を用いて理解を深める点など、可視化の有用性が示されている．しかし、このような Visual Analytics の中でも音楽データに特化した研究事例はまだ少ない．本研究では、音楽データを活用することで Visual Analytics の一つの可能性も示す．

3. 提案手法

本章では、提案手法を構成する各処理の詳細について述べる．図 1 に本手法の流れを示す．

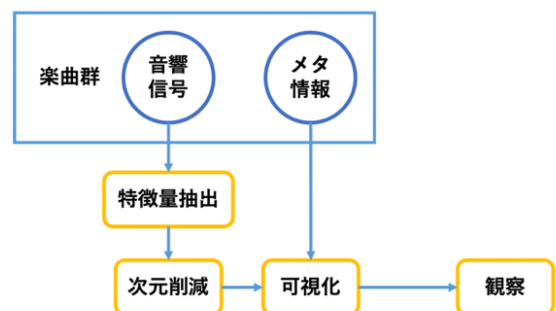


図 1 提案手法の流れ

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。そして得られた音響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。本報告ではこの可視化結果についての観察結果を示す。

3.1 本研究で扱うデータ

本節では、本研究で扱うデータとして楽曲データとメタ情報について説明する。

3.1.1 楽曲データ

本研究では入力データとして各楽曲の音響信号を扱う。本報告で使用している音響データは 1986 年から 2018 年までの日本のヒット曲 1315 曲からなる楽曲群を適用した。楽曲は全て m4a データである。

3.1.2 メタ情報

本報告では、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。本報告では、楽曲群の各楽曲に確実に存在するメタ情報である年代、作曲家、アーティストを可視化に用いた結果を紹介する。楽曲データには各曲の印象をはじめとした主観をとまなうメタ情報が付与されていることがあるが、本報告ではそのようなメタ情報は適用していない。

3.2 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。本研究では以下の 2 種類の音響特徴量

- ・ 音楽的な説明性のある音響特徴量
- ・ 機械学習が生成する音響特徴量

を対象とする。これらの音響特徴量を算出するツールを 1 種類ずつ用いることで、どちらかに偏った結果の分析を避けることや、音楽的説明性のある音響特徴量あるいは機械学習が生成する音響特徴量により適した次元削減手法や可視化手法を探し出すことを目指す。

3.2.1 音楽的な説明性のある音響特徴量

本研究では、楽曲の速度、音量などのように説明的で理解しやすい音響特徴量を「音楽的な説明性のある音響特徴量」と呼ぶ。本研究では、音楽的な説明性のある音響特徴量を Librosa[11]を用いて算出する。Librosa は、さまざまな音響特徴量を抽出し、音楽や音声を分析/解析することができる、Python の MIR コアライブラリである。簡易なプログラムで音響特徴量を抽出できるため、本研究では Librosa を採用した。Librosa で算出できる音響特徴量のうち、本研究で利用しているものを表 1 に示す。これらの音響特徴量によ

って楽曲の特徴を客観的に数値化することで、メタ情報との関係を可視化できると考えて選択した。このうち、rms energy, spectral contrast, spectral flatness, spectral rolloff の 4 種類については、数秒単位フレームごとに算出されるため値が複数次元で算出されるが、本研究ではこれらの平均と分散を計算し、それぞれ単一の値として使用する。

表 1 Librosa で算出する音響特徴量（全て 1 次元）

tempo	楽曲の速度
rms energy	エネルギーの二乗平均平方根
spectral contrast	周波数の明瞭さ
spectral flatness	周波数別エネルギーの均一さ
spectral rolloff	エネルギー85%となる周波数

3.2.2 機械学習が生成する音響特徴量

本研究で採用する「機械学習が生成する音響特徴量」は高次元なベクトルとして算出されるのが一般的である。個々の次元は説明性をもたないが、楽曲間の類似度をベクトル間距離で表現することができる。本研究では、機械学習が生成する音響特徴量を RP_extract[12]を用いて算出する。RP_extract は、ウィーン工科大学独自のモデルで音響特徴量を計算する機械学習ツールであり、機械学習を用いて 3 種類の音響特徴量を算出する。3 種類の音響特徴量について表 2 に示す。リズムのパターンやビートやエネルギーの発生に関する音響特徴量は、Librosa で算出する音響特徴量とは異なる性質を有するものである。これを適用することで、メタ情報と音響特徴量との間の別の関係を発見できることを期待して、本研究では RP_extract を採用した。

表 2 RP_extract で算出する音響特徴量

リズムパターン (RP) 1440 次元	人間の可聴範囲のクリティカルバンドごとの変調振幅(リズム)を表す。
リズムヒストグラム (RH) 60 次元	全てのクリティカルバンドの変調周波数が合計され、変調周波数ごとの“リズムカルなエネルギー”のヒストグラムとして表す。
統計スペクトル記述子 (SSD) 168 次元	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、記述する。

3.3 次元削減

3.2 節で算出した音響特徴量を、Librosa は 5 次元から、RP_extract は 1668 次元から 2 次元に次元削減する。算出する音響特徴量は多次元データとなり、そのまま可視化することは困難である。そこで本研究では、

多次元から2次元への次元削減を適用し、これを散布図として表示することで、2次元平面上での可視化が可能となり、楽曲群への定性的解釈を支援できると考えた。また本研究では、複数の次元削減手法を適用することで、手法による可視化結果の違いや、どの次元削減手法が音楽分類に効果的であるかを検討する。現在本研究で使用している次元削減手法はPCA、MDS(Multi-Dimensional Scaling)、t-SNE、UMAP(Uniform Manifold Approximation and Projection)の4種類である。その中でも本稿では、メタ情報と音響特徴量の関係が顕著に現れたt-SNEを用いた可視化結果を紹介する。

3.4 可視化

3.3節で算出した音響特徴量を、散布図として楽曲の分布を表示する。大量の楽曲群を点群として一画面に配置し、点群間の距離から近い特徴を持つ楽曲群を視覚的に捉えることができるため、散布図は楽曲群の可視化に適していると考えられる。そのため、本研究では次元削減後の可視化手法として散布図を採用した。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、メタ情報と音響特徴量の関係を把握できるようにする。また、散布図の可視化でメタ情報と音響特徴量の間に傾向が現れた場合、その主な要因となる要素を求めることが必要であると考えられる。散布図による可視化結果を詳細に分析するため、表1に示したLibrosaで算出する音響特徴量を個別に棒グラフにして可視化する。このとき、特定の音響特徴量を全楽曲分並べて比較したいため、大量の数値の大きさの差を一度に視覚的に捉えやすくなる棒グラフを採用した。

3.5 観察

3.4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。
- ・ どの音響特徴量や機械学習を用いると明確な傾向を持っ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。

4. 実行例と考察

1986年から2018年までの日本のヒット曲1315曲からなる楽曲群を適用して提案手法を実行した。本章では、LibrosaおよびRP_extractで抽出したそれぞれの音響特徴量において、全特徴量を包括的に可視化した結果と個別に可視化した結果について議論する。また、顕著な特徴が見られた可視化の一部を紹介する。

4.1 Librosaを用いた可視化

Librosaで算出した音響特徴量を用いた可視化では、どのメタ情報と次元削減手法を使っても、現時点でメタ情報と音響特徴量との間に強い関係は見られなかった。そのため散布図の例示は割愛する。

4.2 RP_extractを用いた可視化

RP_extractで算出した音響特徴量を全てまとめてt-SNEで次元削減し、散布図として表示した。可視化結果について考察する。

年代、作曲家、アーティストで色分けした散布図を、それぞれ図2,3,4に示す。年代による色分けでは、全楽曲を6つの年代に分類して色分けした。作曲家による色分けでは、データ中に20曲以上が存在する4人の作曲家に対して色を割り当てた。また、アーティストによる色分けでは、データ中に18曲以上が存在する10組のアーティストに対して色を割り当てた。いずれの場合にも、色の割り当てに該当しない楽曲は「その他」として同一色で可視化した。

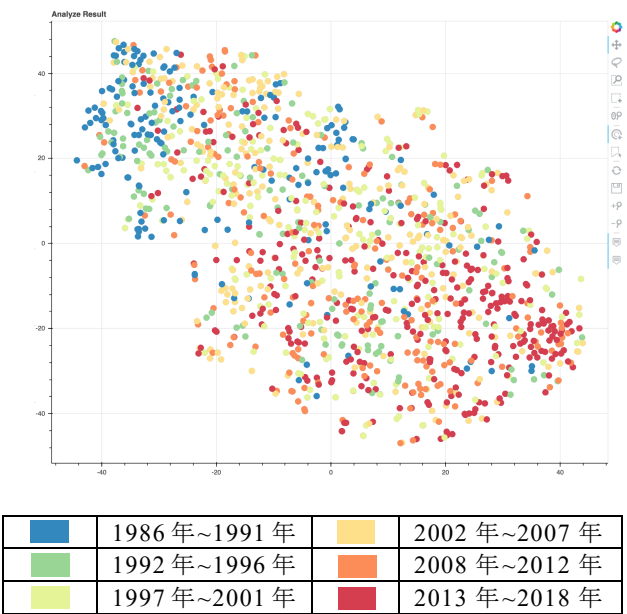
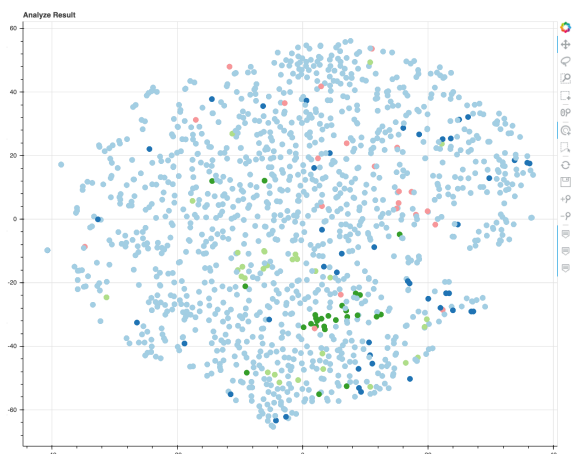
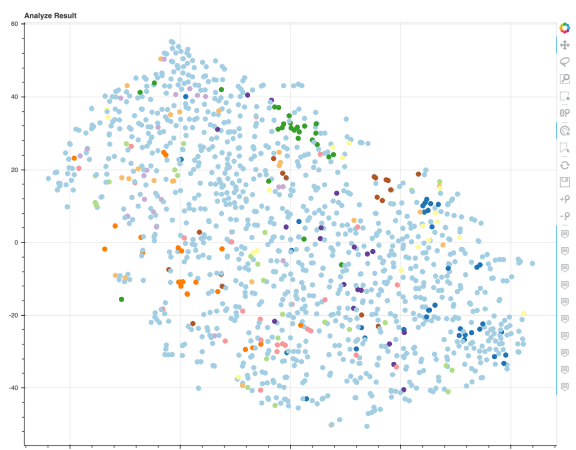


図2 年代で色分けした可視化結果 (RP_extractの全音響特徴量を次元削減)



■	小室哲哉	■	織田哲郎
■	後藤次利	■	その他
■	桜井和寿		

図3 作曲家で色分けした可視化結果(RP_extractの全音響特徴量を次元削減)



■	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		

図4 アーティストで色分けした可視化結果(RP_extractの全音響特徴量を次元削減)

年代で色分けした可視化結果(図2)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。作曲家で色分けした可視化結果(図3)では、桜井和寿氏(緑)の楽曲が一点に集中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(図4)では、Mr.Children(緑)に特に明確なクラスターが見られる。その他、ZARD(茶)、AKB48(青)、EXILE(オレンジ)、ゆず(紫)などの一部の楽曲もクラスターを形成していることが見受けられる。

4.3 RP_extractによる各音響特徴量を用いた可視化
4.2節で示した可視化について、メタ情報と音響特徴量のさらに具体的な関係を調べるため、表2で示したRP_extractの3種類の音響特徴量を分割し、それぞれについて同様にt-SNEを適用して可視化した。すると図5, 6, 7に示すように、SSDのみを利用したときに、全ての音響特徴量を使用した時と同じような傾向が現れることが示された。これによって、1986年から2018年までの楽曲間や、作曲家ごと、アーティストごとの楽曲間では、リズムよりも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。

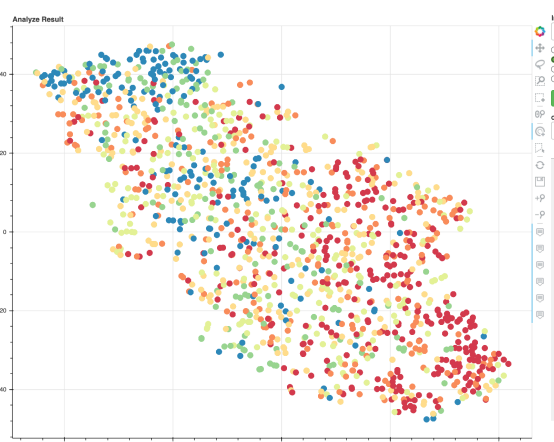


図5 年代で色分けした可視化結果(RP_extractよりSSDのみを次元削減, 凡例は図2と同様)

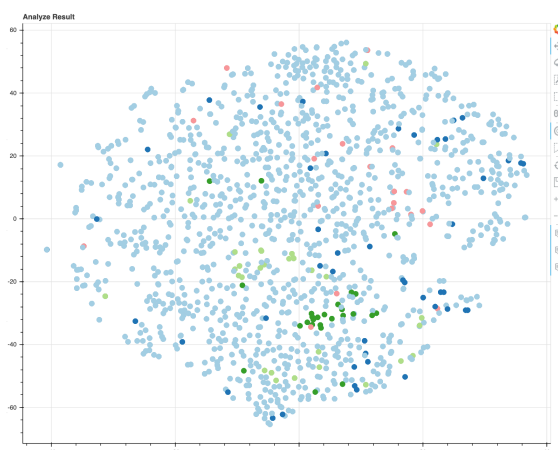


図6 作曲家で色分けした可視化結果(RP_extractよりSSDのみを次元削減, 凡例は図3と同様)

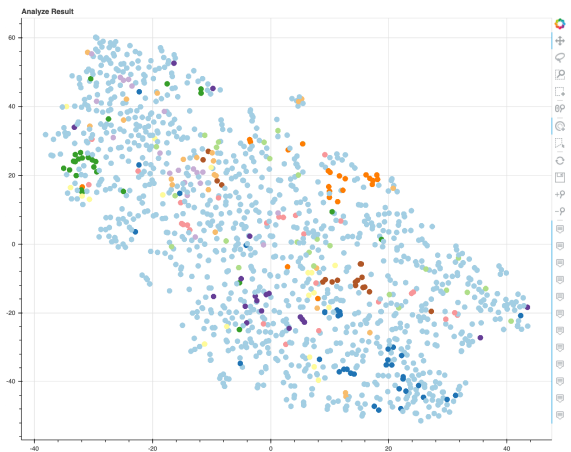


図 7 アーティストで色分けした可視化結果 (RP_extract より SSD のみを次元削減, 凡例は図 4 と同様)

4.4 Librosa による各音響特徴量を用いた可視化

4.3 節にて示した図 5, 6, 7 の可視化より, 年代間, 作曲家間, アーティスト間で楽曲のビートやエネルギーに特徴の違いがあることを発見した. 実際にどのような違いがあるかを確認するため, Librosa で算出した音響特徴量を個別に可視化した. 以下の可視化結果は全て, 楽曲を左右に並べて, 所定の音響特徴量を棒グラフで表示したものである.

図 8 は rms energy の平均値を縦軸として, 年代による色分け (図 2 参照) を適用した可視化結果であり, 6 つの年代順に楽曲を左右に並べている. 図 9 は spectral flatness の平均値を縦軸として, 作曲家による色分け (図 3 参照) を適用した可視化結果であり, 作曲者別に楽曲を左右に並べている. 図 10 は spectral flatness の分散を縦軸として, アーティストによる色分け (図 4 参照) を適用した可視化結果であり, アーティスト別に楽曲を左右に並べている. 図 9, 図 10 については「その他」に属する楽曲数が多いため, 「その他」に属する楽曲を除いた結果もそれぞれ図 11, 12 に示している.

rms energy の平均は概ね年代が上がるほど大きくなっている. rms energy は音量に相当するエネルギーの二乗平均平方根である. エネルギーの最大値は過去の楽曲も最近の楽曲もほぼ同一であることから, 過去の楽曲ほどエネルギーの変化に富み, 最近の楽曲ほどエネルギーが均一化されていることを示す.

spectral flatness の平均は, 他の作曲家に比べて桜井和寿氏(緑)が特に低くなっている. 桜井和寿氏の楽曲は周波数別のエネルギーの均一さが低いことが示された. また, 桜井和寿氏の作曲した楽曲は全て

Mr.Children の楽曲であるため, アーティストごとの色分けでも顕著な結果を示す一端となった.

spectral flatness の分散は, Mr.Children(緑)だけでなく, 浜崎あゆみ(薄オレンジ)や EXILE(オレンジ)のような一部のアーティストの値が低くなっている. これらのアーティストの楽曲は spectral flatness の平均も比較的低く, 周波数別のエネルギーの均一さが曲を通して低いという結果となった.

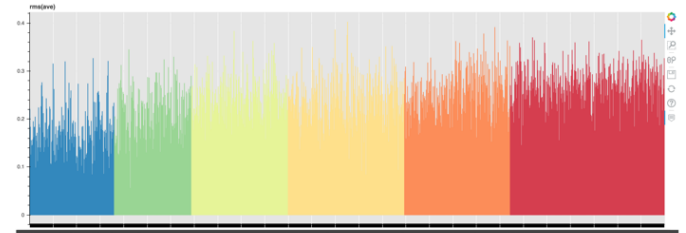


図 8 年代で色分けした可視化結果(横軸は年代順, 縦軸はその曲の rms energy の平均値, 凡例は図 2 と同様)

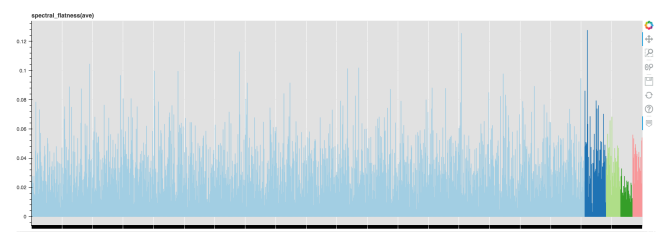


図 9 作曲家で色分けした可視化結果(横軸は作曲家ごと, 縦軸はその曲の spectral flatness の平均値, 凡例は図 3 と同様)

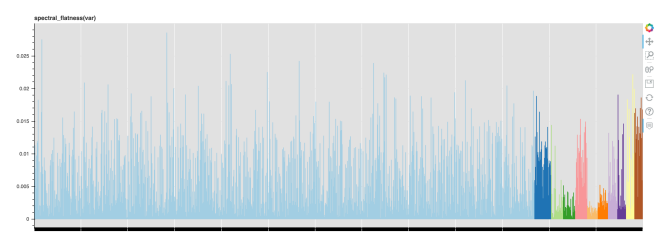


図 10 アーティストで色分けした可視化結果(横軸は各楽曲, 縦軸はその曲の spectral flatness の分散値, 凡例は図 4 と同様)

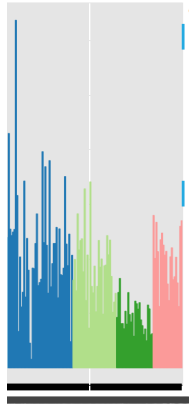


図 11 作曲家で色分けした可視化結果のうち、その他のアーティストを除いたもの(横軸は作曲家ごと、縦軸はその曲の spectral flatness の平均値、凡例は図 3 と同様)

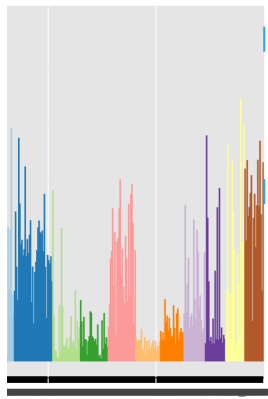


図 12 アーティストで色分けした可視化結果のうち、その他のアーティストを除いたもの(横軸は各楽曲、縦軸はその曲の spectral flatness の分散値、凡例は図 4 と同様)

4.5 考察

以上の結果から、Librosa で算出した音響特徴量をまとめて次元削減した場合には、年代、作曲家、アーティストによる楽曲の特徴は発見できなかったのに対して、RP_extract で算出した音響特徴量を t-SNE で次元削減した場合にメタ情報と音響特徴量の間に特定の傾向が現れることを発見した。今回は t-SNE による次元削減結果のみを例示したが、適用した 4 種類の次元削減手法では、PCA や MDS よりも t-SNE や UMAP でメタ情報と音響特徴量の間に傾向が現れていた。この要因として、PCA と MDS が線形手法であるのに対し、t-SNE と UMAP は非線形手法であることが考えられる。非線形手法では大域的な特徴と局所的な特徴をどちらも保持しやすく、今回の音響特徴量のデータにおいても適切な手法であったと言える。

また、具体的に、メタ情報と音響特徴量の関係とし

て、年代間では rms energy の平均値に、作曲家間では spectral flatness の平均値に、アーティスト間では spectral flatness の分散値に特徴があるということがわかった。rms energy の平均値は概ね年代が上がるほど大きくなっている。これは過去の楽曲ほどエネルギーの変化に富み、最近の楽曲ほどエネルギーが均一化されていることを示している。この主な原因として、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。spectral flatness の平均値は、他の作曲家に比べて桜井和寿氏(緑)が特に低くなっており、桜井和寿氏の楽曲は周波数別のエネルギーの均一さが低いことが示された。spectral flatness の分散値は、Mr.Children(緑)だけでなく、浜崎あゆみ(薄オレンジ)や EXILE(オレンジ)のような一部のアーティストの値が低くなっており、これらのアーティストの楽曲はエネルギーの均一さが曲を通して低いことが示された。これにより、これらのアーティストには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。

本研究では、可視化を通して音響特徴量とメタ情報の新たな関係を発見できるかを実証することを目的の一つとしていたが、上記の結果より、これは実証できたと言える。

また、これらは今回使用した音響特徴量や楽曲群、メタ情報によるものであり、他の音響特徴量や楽曲群、メタ情報を使用した際にはこの限りではないと考えられる。

5. まとめと今後の課題

本研究では、音楽分類タスクに効果的なメタ情報、音響特徴量、機械学習手法、可視化手法を検討すること、また音響特徴量とメタ情報の新たな関係を可視化によって発見できるか否かを実証することを目標としている。その一環で我々は、メタ情報を付与した楽曲群について、音楽解析ツールや機械学習を用いて音響特徴量を算出し、メタ情報と音響特徴量の分布を可視化してその結果の観察した。本報告ではその進捗として、可視化結果の一部を紹介した。その中でも特に、RP_extract で算出した音響特徴量を t-SNE で次元削減して散布図にすることで、メタ情報と音響特徴量の間に特定の傾向が現れることを発見した。また、RP_extract の中でも SSD が年代や作曲家、アーティストによる特徴を顕著に示すことがわかった。具体的には、1980 年代から 2010 年代の各時代の楽曲には、時代の流れとともにビートやエネルギーに変化が見られることがわかった。また、特定の作曲家・アーティストについて、ビートやエネルギーの特徴が類似する曲が多数見られることがわかった。さらに、年代間では

rms energy に、作曲家間および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。

本研究より、可視化を通して音響特徴量とメタ情報の新たな関係の発見を実証することができた。今回扱った楽曲群の音響特徴量では、利用した次元削減手法のうち、非線形手法の1つである t-SNE が最も適していると考えられた。t-SNE で次元削減し、散布図にするという手法は、音響特徴量に限らず、多次元データとメタ情報の特定の関係を把握することに役立つ可能性があるかと推察される。また、扱うメタ情報やデータによって適切な次元削減手法や可視化手法が異なる可能性もあるが、本手法のように複数の可視化結果を比較することが、適切な要素を特定するのに有用であると考えられる。

今後の課題として、実験に扱ったメタ情報、音響特徴量や可視化手法のさらなる検討があげられる。特に、図 5~7 の可視化結果について、ビートやエネルギーに特徴があること、および図 9~12 の可視化結果について、各メタ情報と音楽的に意味のある音響特徴量の関係を論じているが、Librosa が算出する音響特徴量のうち表 1 に示した 5 種類のみではその特徴を十分に説明できないと考えている。そこで、今回採用した 5 種類以外の音響特徴量を算出した結果を観察することで、RP_extract による可視化結果に関する考察をさらに深めたい。さらに、本研究に特化した音響特徴の機械学習モデルを構築して適用したい。本報告にてメタ情報と音響特徴量の新たな関係を表現できた t-SNE は、調整可能なパラメータ perplexity を有する。本研究では初期値の一定のまま全ての実験をおこなったが、パラメータを適切に調整することでデータの大域的な特徴と局所的な特徴のどちらを優先するかのバランスを調整でき、違った可視化結果が得られることを期待できる。また、今回の実験で興味深い可視化結果があまり見られなかった PCA, MDS, UMAP についても、それに適した音響特徴量が存在しないか検討する余地がある。複数の次元削減を組み合わせることによって、さらに明確にクラスがわかれる可能性もあるため、これについても試みる必要がある。そして、これらについて検討や実装しつつ、音楽製作者・販売者が使用することも想定し、操作性も追求したシステムとして実装を進めたい。

参 考 文 献

- [1] 一般社団法人日本レコード協会, 音楽メディアユーザー実態調査 2019 年度調査結果報告書 公表版. <https://www.riaj.or.jp/f/pdf/report/mediauser/softuser2019.pdf>, 2020.
- [2] Spotify, <https://www.spotify.com> (2022-02-08 参照).
- [3] Keunwoo Choi, Gyorgy Fazekas, Mark Sandler and Kyunghyun Cho, "Transfer learning for music classification and regression tasks," 18th International Society of Music Information Retrieval (ISMIR) Conference, 2017.
- [4] Wei Li and Jiake Li, "Research on Music Visualization Based on Graphic Images and Mathematical Statistics," IEEE Access, 2020.
- [5] Juan Jose Burred, Alexander Lerch, "A Hierarchical Approach To Automatic Musical Genre Classification," Proc. of the 6th Int. Conference on Digital Audio Effects (DAFx-03), London, UK, September 8-11, 2003.
- [6] Kin Wai Cheuk, Yin-Jyun Luo, Balamurali B, T, Gemma Roig and Dorian Herremans, "Regression-based music emotion prediction using triplet neural networks," IJCNN 2020, 2020.
- [7] Jiyoung Park, Jongpil Lee, Jangyeon Park, Jung-Woo Ha and Juhan Nam, "Representation Learning Of Music Using Artist Labels," International Conference on Machine Learning (ICML) 2019, 2019.
- [8] 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, At Limassol, Cyprus, 2018.
- [9] Arthur Flexer, "Improving Visualization of High-Dimensional Music Similarity Spaces," 16th International Society for Music Information Retrieval, 2015.
- [10] Wilson Estécio Marcílio Júnior, Danilo Medeiros Eler and Rogério Eduardo Garcia, "Contrastive analysis for scatterplot-based representations of dimensionality reduction," Computers & Graphics, Volume 101, December 2021, pp. 46-58 .
- [11] Librosa development team, Librosa, <https://librosa.org/doc/latest/index.html#> (2022-02-08 参照).
- [12] Vienna University of Technology, Music Information Retrieval, <http://ifs.tuwien.ac.at/mir/musicbricks/#RPextract> (2022-02-08 参照).