

日本の音楽の人気度分析

竹之内香穂 (251x040x)

1. はじめに

現代の音楽市場において、楽曲の人気要因を理解することは、アーティストやプロデューサーといった制作側のみならず、大衆文化の変遷をたどる研究者にとっても重要な課題となっている。しかしながら、膨大な楽曲が供給され続ける現代のストリーミング環境において、多種多様な楽曲の中から共通の特徴を特定することは困難である。そこで本研究では、人気な楽曲にはどのような特徴があるのか特定することを目的とする。

音楽心理学は、人間が音楽を通じてどのように感情が変化するかを解明するための分野である。この分野において、楽曲の感情的特性を定量化する際、最も基礎的な理論の一つとされるのが「感情の円環モデル」[1]である。これは、あらゆる感情は「覚醒-沈静(arousal)」と「快-不快(valence)」の2軸で説明できるというものである。近年の音楽業界はストリーミングサービスによって大きく変容し、楽曲の特徴に関する大規模なデータセットが入手可能となった。その中でも、Spotify等の配信プラットフォームから提供される energy の指標は、楽曲の知覚的強度を示すものであり、感情の円環モデルにおける「覚醒-沈静(arousal)」に対応する指標と位置づけられる。したがって本研究では、Spotify 楽曲データセットを用いて、popularity, energy, valence といった特徴量間の統計的関係を分析するためのシステムを提案する。

2. 手法

本研究では、HTML, CSS, D3.js(v7)を用いて分析システムを構築した。本システムは、energy と valence の分布可視化(散布図)、統計情報ビュー、ジャンル構成(棒グラフ)から構成される(図 1)。

まず、メインビューとなる散布図(図 1 左)では、縦軸に energy、横軸に valence をとり、各楽曲の分布を可視化した。各点の色は popularity を表している。インタラクション機能として、散布図上をドラッグすることで任意の範囲を選択可能とした。これにより、選択した範囲に含まれる楽曲のジャンル構成(棒グラフ)と統計情報ビューが動的に更新される。デフォルトでは、散布図に表示されている全楽曲が選択されている。また、画面上部の Minimum Popularity バーを操作することで、一定の人気度以上の楽曲のみを表示するフィルタリングが可能である。個々の楽曲については、点にカーソルを合わせることで、その点が示す楽曲のタイトル・アーティスト名・人気度が表示される機能を実装した。さらにクリックすると、その楽曲のジャケット画像と Spotify での視聴リンクが統計情報ビューに表示される機能を実装した。

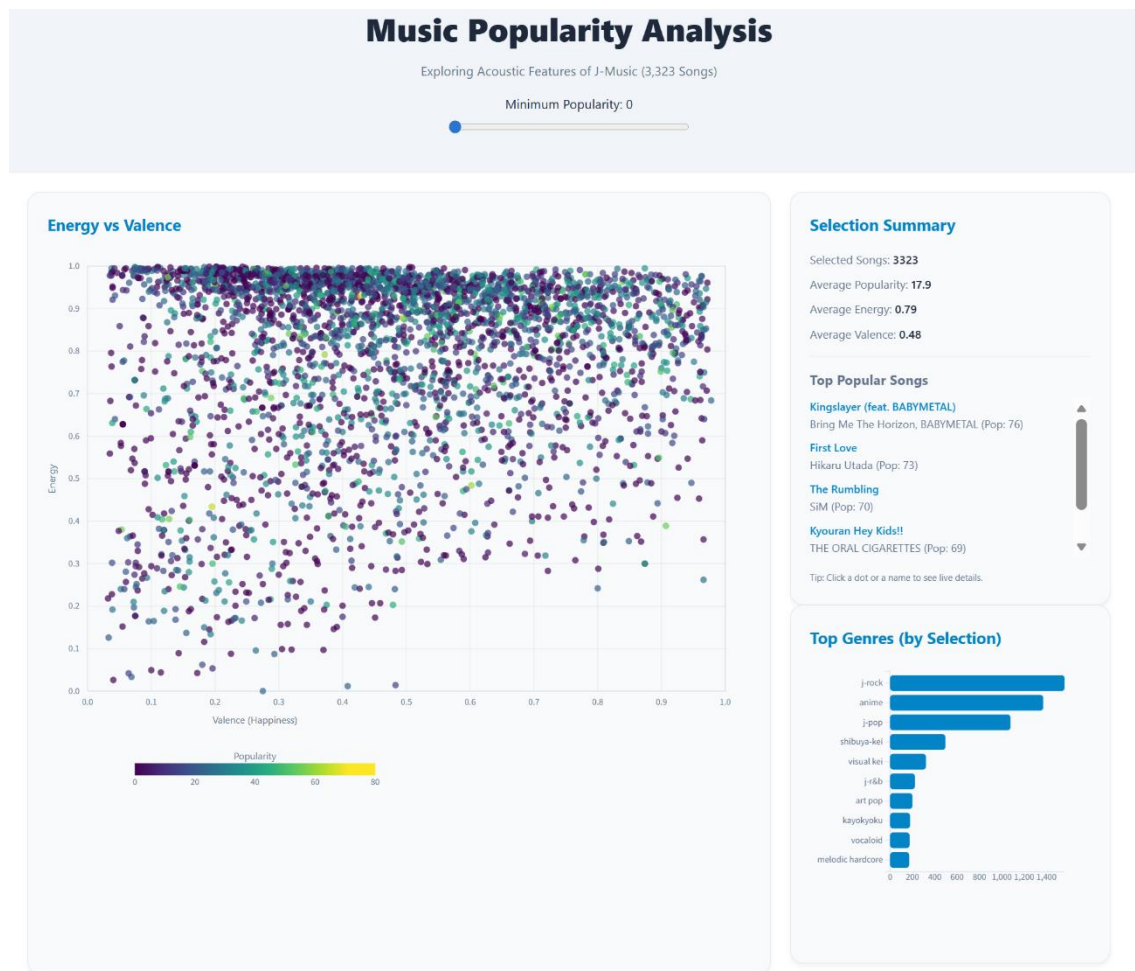


図 1：システム概観

次に、詳細ビュー(図 1 右上)では、選択した範囲内における楽曲数および popularity, energy, valence の平均値を表示する。さらに、選択範囲内において popularity が高い、すなわち人気度が高い上位 5 曲のリストを表示することで、注目する楽曲の具体的な特徴量の値を確認できるようにした。最後に、棒グラフ(図 1 右下)では、選択した範囲に含まれる楽曲のニッチジャンルの分布を表示し、当該領域においてどのようなジャンルの楽曲が多いかを可視化している。

3. 実験

本研究では、Kaggle 上の Spotify の約 55 万曲のデータセット [3] から、日本の楽曲（ニッチジャンルに “j-pop”, “anime”, “j-rock”, “japanese”, “vocaloid”, “shibuya-kei” が含まれている楽曲）3323 曲を抽出したデータセットを用いた。実験では、まず全楽曲を対象として、使用データ全体の特徴を分析した。次に、popularity が高い楽曲を対象として、人気な楽曲の特徴を分析した。

3.1 全楽曲に対する分析

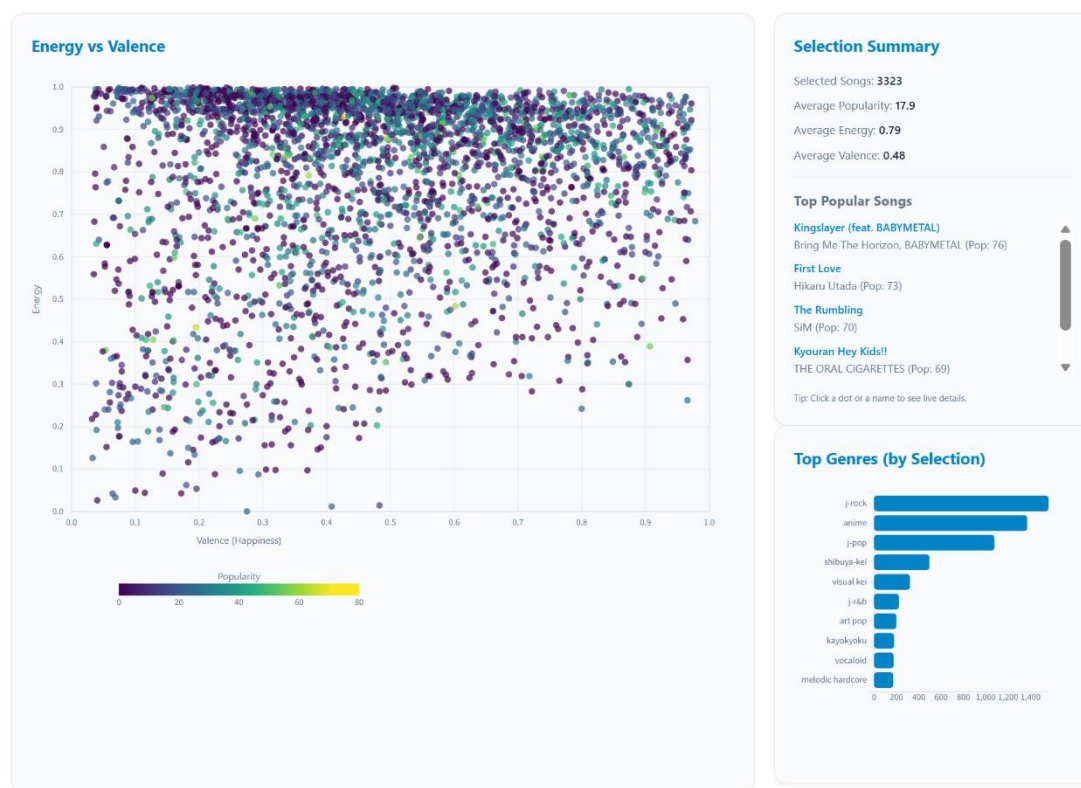


図 2：全楽曲を選択したときの結果

使用データにおける全 3323 曲に対して分析を行った(図 2)。まず散布図を見ると、上部に点が多くプロットされていることが確認できる。したがって、本データセットにおいては、valence の値にかかわらず energy の値が大きい楽曲が支配的であることがわかる。一方で、散布図の右下(energy が約 0.2 以下かつ valence が約 0.5 以上)の部分には点がプロットされていないことが確認できる。したがって、静かで明るい楽曲は本データセットには存在しないことがわかる。また、点の色に注目すると、紫色や濃い緑色の点が多く、黄色や黄緑色の点が少ないことが確認できる。したがって、本データセットの大半が popularity の値が小さい楽曲で構成されていることがわかる。次に詳細ビューを見ると、energy, valence の平均値はそれぞれ 0.79, 0.48 であることが確認できる。最後に棒グラフを見ると、ジャンル構成は”j-rock”が最も多く、次いで”anime”、”j-pop”の順であることが確認できる。

3.2 人気な楽曲に対する分析



図3：人気曲を選択したときの結果

Minimum Popularity バーを調節し(Minimum Popularity:48)、全 3323 曲の中で popularity が高い 213 曲を抽出して分析を行った(図 3)。まず散布図を見ると、図 2 と同様に上部に点が多くプロットされていることが確認できる。したがって、valence の値にかかわらず energy の値が大きい傾向は維持されていることがわかる。一方で、図 2 との明確な差異として、散布図の左上(energy が約 0.7 以上かつ valence が約 0.1 以下)および左下(energy が約 0.2 以下かつ valence が約 0.5 以下)の部分に点がプロットされていないことが確認できる。したがって、人気楽曲においては、激しくて暗い曲や静かで暗い曲は存在しないことがわかる。次に詳細ビューを見ると、energy, valence の平均値はそれぞれ 0.81, 0.54 であることが確認できる。これは図 2 の全楽曲平均と比べると、どちらも値が大きくなっていることが確認できる。最後に棒グラフを見ると、ジャンル構成は”j-pop”が最も多く、次いで”anime”、”j-rock”の順であることが確認できる。これは図 2 のときと比べると、人気曲においては”j-rock”よりも”j-pop”の割合が高いことがわかる。

4. 考察

本実験の結果から、使用したデータ全体と比較して、人気度が高い楽曲の特徴について考察を行う。全体的な傾向・全楽曲と人気曲との差異・energy と valence の平均値の差異・ジャンル構成の 4 つの観点で考察を行う。

第一に、全体的な傾向について考察を行う。全楽曲(3.1 節)と人気曲(3.2 節)の双方において、valence の値にかかわらず energy の値が大きい楽曲が多かった。したがって、本実験で使用するデータセットにおいては、popularity や valence の値にかかわらず、energy の値が大きい、すなわち激しい楽曲が多い傾向があると考えられる。

第二に、全楽曲と人気曲との差異について考察を行う。全楽曲(3.1 節)と人気曲(3.2 節)の分布の差として、人気楽曲の中には激しくて暗い曲や静かで暗い曲は存在しなかった。したがって、日本の人気楽曲においては、激しさ(energy)にかかわらず、ある程度の明るさ(valence)が重要であり、暗い楽曲は人気が出にくいと考えられる。

第三に、平均値の差異について考察を行う。人気曲(3.2 節)における energy, valence の平均値は、全楽曲(3.1 節)と比較してそれぞれ大きくなっていった。したがって、全体と比較して人気曲はより激しく明るい曲調の傾向があると考えられる。

最後に、ジャンル構成について考察を行う。全楽曲(3.1 節)では”j-rock”が最多であったのに対し、人気曲(3.2 節)では”j-rock”よりも”j-pop”の方が多かった。これは、特定のファン層を持つロックよりも、より大衆向けであるポップスの方が人気を得られやすいことを示唆していると考えられる。

5. まとめ

本研究では、日本の人気楽曲にはどのような特徴があるのかを、popularity, energy, valence といった特徴量を用いて可視化・分析するシステムを構築した。energy と valence の分布可視化(散布図)、統計情報ビュー、ジャンル構成(棒グラフ)を連動させることで、楽曲データの対話的な分析を可能にした。

Spotify 楽曲データセットから抽出した日本の楽曲 3323 曲を対象とした実験の結果、データセット全体として、valence の値にかかわらず energy の値が大きい楽曲が支配的であることが確認された。さらに、popularity が高い楽曲に限定して比較分析を行ったところ、全楽曲平均よりも energy, valence の平均値がそれぞれ高く、特に valence が低い楽曲は存在しないという特徴が明らかになった。また、ジャンル構成においては、人気楽曲では”j-rock”よりも”j-pop”の割合が高いことが示された。

最後に、本研究における課題とそれに基づく今後の展望について述べる。第一に、データセットに関する課題である。本実験では公開データセットから特定のニッチジャンル(”j-pop”, ”anime”等)を含む楽曲のみを抽出して分析を行った。しかしながら、これは日本の音楽市場全体を完全に網羅しているとは限らない。今後は Spotify Web API 等を用いて、より幅広いデータを取得・分析することで、結果の信頼性を高める必要がある。第二に、可視化手法に関する課題である。実験結果(図 2, 図 3)を見ると、散布図の上部に点が密集していることが確認できる。これにより、データが密集している領域において分析が困難であった。今後は点の透明度を調整するだけでなく、密度等高線やヒートマップなどを導入することで、より正確に可視化する必要がある。最後に、多角的な分析に関する課題である。本研究

では主に popularity, energy, valence の関係に注目したが、楽曲の人気要因はこれらのみに留まらない。他の特徴量を用いた分析は、今後の課題である。

6. 参考文献

[1] Russell, J. A. (1980). "A circumplex model of affect ". *Journal of Personality and Social Psychology*.

[2] Kaggle Dataset : 550K Spotify Songs: Audio, Lyrics & Genres

<https://www.kaggle.com/datasets/serkantysz/550k-spotify-songs-audio-lyrics-and-genres>

・ Munzner, Tamara. (2014). "Visualization Analysis and Design." *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Courses*: