

状況認識を活用した楽曲プレイリストの自動生成 Automatic Playlist Generation Using Context Awareness

内山 稜凱

Ryoga Uchiyama

法政大学情報科学部ディジタルメディア学科

E-mail: ryoga.uchiyoama.4k@stu.hosei.ac.jp

Abstract

The spread of the Internet and music subscription services has enabled us to listen to songs anywhere with smartphones. This also allows us to listen to songs that are suitable for various situations. However, it is still inconvenient to generate playlists of songs. These services only recommend songs with high preferences and are not able to play songs that are suitable for current contexts. This paper proposes a system that recognizes the current context and automatically generates a playlist of songs that matches the context. The system uses five features, danceability, loudness, energy, tempo, and valence as indices for music selection. For this system, we analyzed a total of 300 songs and selected about 100 songs with the highest or lowest features. Using the selected songs, the system automatically generates a playlist of 10 songs that is appropriate for the context. It uses four types of contexts (i.e., weather, current location, action, and emotion), each of which consists of four instances (e.g., joy, anger, sadness, and pleasure for emotion). In the experiment conducted to evaluate the system, participants were asked to score the songs in the generated playlists to see if they were appropriate for the contexts, which resulted in relatively high scores on average.

1. はじめに

音楽は太古より我々を楽しませてくれるエンターテインメントの 1 種である。近年ではインターネットやサブスクリプションサービスの普及により、スマートフォン 1 つさえあればどこでも音楽を鑑賞することができるようになった。これにより音楽を鑑賞できる状況、環境が増え、それに適した楽曲を聞くことが増えた。しかし、楽曲のプレイリストの生成は不便である。プレイリストを生成する際に、再生する状況を想定して 1 曲 1 曲手動で追加しなければならない。

こうした不便を解決する手段として、Spotify Mixes や Apple Music Favorites Mix 等の自動プレイリスト生成サービスが提供されているが、嗜好度の高い楽曲を推薦するだけで、実際に楽曲を再生する際の情景や感情に寄り添った楽曲を再生できない。風景画像から状況を推定し、関連の高い楽曲を推薦するシステム [1] も実現され、状況を活用した楽曲推薦、プレイリスト生成を可能にしてい

るが、使用する際に風景画を入力しなければならない点、リアルタイムでの使用が難しい点などの問題がある。

本研究では、楽曲を再生する状況を認識し、情景や感情に適した楽曲を組み込んだプレイリストを自動生成するシステムを提案する。認識する状況は、現在地、天気、行動、感情の 4 種類とし、GPS やユーザーからの入力などをもとにデータを取得する。楽曲の選択には、Danceability, Energy, Loudness, Tempo, Valence の 5 種類の音楽特微量を用い、300 曲から最終的に 10 曲を選択してプレイリストを作成する。本システムを用いて 14～22 歳の男女 8 人を対象に実験を行った結果、5 ポイント中平均 3.27 ポイントの評価を得た。しかし、適切な楽曲が選択できない状況があることもわかった。

2. 関連研究

桐本ら [1] はユーザーの入力した風景画像の色彩情報から得られる環境状況と印象との相関を計測することで、入力されたデータとの相関の高い楽曲を推薦するシステムを実現した。この研究では入力された風景画像から天気を推定し、そこから感性データを取得する。このデータと実際のユーザーの感性情報を照らし合わせることで関連の高い音楽データを抽出した。

Spotify によるプレイリストサービス Spotify Mixes は、ユーザーの聴取傾向にパーソナライズしたプレイリストを自動的に生成するもので、Spotify で楽曲を聴けば聞くほど聴取傾向や嗜好が反映される。

感情と音楽の関係について数多くの研究が存在する。大出ら [2] は、音楽聴取時の感動評価尺度に差が生まれることから、感動の度合いによる音楽や音響の印象について調査を行った。その結果「音色がよい」や「艶がある」といった音響の印象によって感動の度合いが変化したと結論付けている。

場所や天気等の環境と音楽の関係性についても研究が行われている。加藤ら [3] は、春の童謡に注目し、気象と音楽を連携させた授業の実践を試みている。生徒らに 4 月初めの急昇温について把握させた後、歌唱表現学習を行う実験を行い、気象と音楽の連携の可能性を見出すことに成功している。

3. 提案手法

3.1. 状況認識

本研究では、現在地、天気、行動、感情の 4 種類の状況について認識を行う。現在地、天気についてはスマー

トフォンの機能を用いる。現在地の取得は GPS を用いて緯度経度を取得することで実現し、またそれによって得られた緯度と経度を用いて天気を取得する。天気の取得には OpenWeatherMap API を使用する。取得する情報は、晴れ、曇り等の簡単な情報とする。行動、感情については選択式の質問に答えてもらうことにより取得する。行動についての選択肢は移動中、運動中、作業中、休憩中の 4 つである。また、感情についての選択肢は、喜、怒、哀、楽とする。

3.2. 楽曲選択

生成するプレイリストは 10 曲で構成する。プレイリストに使用する楽曲に偏りが無いよう、システムに使用する楽曲データとして事前に 300 曲から約 100 曲を選択する。

楽曲選択には音楽特徴量を使用する。音楽特徴量は Spotify Web API から以下の 5 つを取得する。

- **Danceability** : 踊りやすさを示す数値。数値は 0～1 で、大きくなるほど踊りやすさが高い。
- **Energy** : 楽曲の過激さを表す数値。速い楽曲や、うるさく騒がしい楽曲は数値が大きくなる。
- **Loudness** : 音量・音圧(db)の平均値
- **Tempo** : 1 分間の拍数
- **Valence** : 楽曲の明るさを示す数値。値が高いほどポジティブに聞こえ、低いとネガティブに聞こえる。

取得する特徴量をこの 5 つとしたのは、ほかの特徴量に比べ楽曲のイメージがしやすく、状況認識機能との組み合わせがしやすいと判断したためである。各特徴量の数値が高い楽曲 5 曲と低い楽曲 5 曲を選択しシステムに使用する楽曲とする。この際、重複した楽曲は除く。

Spotify Web API の Python ライブラリ Spotipy を用いて約 300 曲の特徴量を取得した。上に述べた方法でプレイリスト生成に使用する楽曲として 96 曲を選択した。全体の 300 曲と、選択された 96 曲の特徴量を比較した平行座標グラフを図 1 に示す。特徴量の最大値と最小値を持つ楽曲を選択することで偏りなく楽曲を選択できていることが読み取れる。

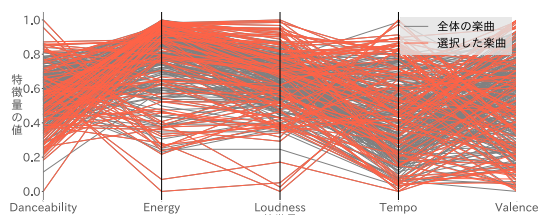


図 1 全体の楽曲と選択した楽曲の特徴量の比較

3.3. プレイリスト生成

事前に選択された楽曲から、状況認識と特徴量を用いてプレイリストのための 10 曲を選択する。10 曲の選択の指標として 2 つの特徴量をピックアップし、それらを計算した値を使用する。

どの状況下でどの 2 つの特徴量を使用するかをアンケート調査によって決定する。状況認識機能で得る状況でどのような楽曲を聴きたくなるかを、特徴量に沿った指標に対して 3 段階評価してもらうことで判断する。特徴量に沿った指標は「踊りやすさ」「過激さ」「音圧」「テンポ」「明るさ」とし、それぞれ Danceability, Energy, Loudness, Tempo, Valence に基づく。また、例として曲名も同時に提示する。アンケート調査の結果によって選択された音楽特徴量を用いて、状況に適した 10 曲のプレイリストを生成して再生する。例えば晴れの時に「明るさ」の数値が高い楽曲が望まれたとすると、実際に認識した天気が晴れの時、曲の明るさを表す Valence の数値が高い楽曲をプレイリストに使用する。

4. 楽曲の特徴量の分析と使用

4.1. アンケート結果の分析

アンケートは 18～24 歳の男女 26 人に解答してもらった。アンケート結果について特に顕著に偏りが見られた項目を図 2 に示す。これは移動中の状況でどのような特徴量を持つ楽曲が好ましいか調査した結果である。評価は 3 段階で行っており、1 が特徴量の数値の低い楽曲、3 が数値の高い楽曲である。この場合、数値の高い楽曲を聞きたいと回答した方が圧倒的に多く、この状況では Danceability の高い楽曲を推薦することを決定している。また、Loudness でも同様に、数値が高い楽曲が好まれる傾向にあることが読み取れる。この状況では Danceability, Loudness の関連性が高いと推測されるため、これらの積を用いることで楽曲の推薦を行う。

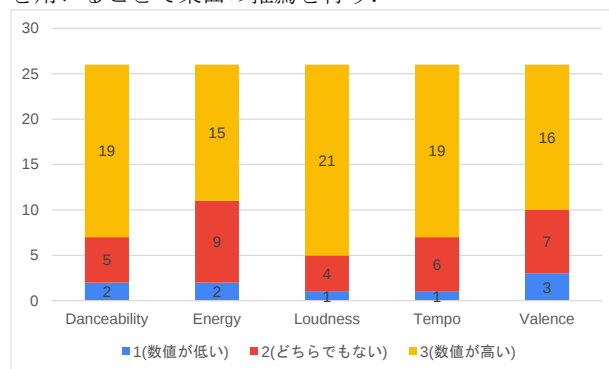


図 2 移動中に聞きたい楽曲の特徴量アンケート結果

4.2. 特徴量の使用

各状況で使用する特徴量の一覧を表 1 に示す。ここでは、Danceability の値を d , Loudness の値を l , Energy の値を e , Tempo の値を t , Valence の値を v として表記する。また、Loudness は -60db から 60db という範囲で数値を取得するため、正の数に計算し直してから使用する。

場所は、表 1 に記載した 4 種類に限定した。当初は GPS によって現在地の位置情報を取得したのち、その周辺施設のデータを得ることによって楽曲選択に役立てる予定だった。しかし、任意の場所で周辺施設のデータを取得することは困難であったため、実験を行う法政大学

小金井キャンパス付近の 4 種類の周辺施設のみを扱うこととした。

表 1 各状況における特徴量の組み合わせ一覧

場所	法政大学小金井キャンパス	$e \times v$ の最小値
	東小金井駅	$(l + 60) \times e$ の最大値
	小金井公園	$t \times e$ の最小値
	その他	$d \times v$ の最小値
天気	晴れ	$e \times v$ の最大値
	曇り	$(l + 60) \times e$ の最小値
	雨	$t \times (l + 60)$ の最小値
	その他	$d \times v$ の最大値
行動	移動中	$d \times (l + 60)$ の最大値
	運動中	$t \times d$ の最小値
	作業中	$t \times v$ の最大値
	休憩中	$t \times d$ の最大値
感情	喜	$(l + 60) \times v$ の最大値
	怒	$e \times v$ の最小値
	哀	$d \times e$ の最小値
	楽	$d \times e$ の最大値

5. 実装

5.1. 状況認識

音楽プレイヤーの画面に遷移する前に、質問によって現在の行動の情報を取得するための画面を図 3(a)に示す。感情選択の際にも類似の画面が表示される。現在地は GPS によって緯度と経度を取得し、その値を使用して周辺施設の情報を取得している。また、GPS によって得られた緯度と経度を OpenWeatherMap API で使用することで現在地の詳細な天気を取得する。本システムで取得する天気情報は、楽曲選択に活用しやすいように、簡単な情報だけを取得している。すべてのデータを取得したのち、データ送信ボタンを押すことで音楽プレイヤーの画面に遷移し、自動生成されたプレイリストが再生される。

5.2. 音楽プレイヤーの作成

自動生成したプレイリストを再生する機能も実装する。プレイヤー自体はシンプルなものを作成する。機能として、再生、次の楽曲、前の楽曲、シークバーによる再生位置の調節、シークバーによる音量の調節、ネットワークの表示、プレイリストの一覧表示を実装している。音楽プレイヤーの画面に遷移した状態の画面を図 3(b)に示す。この画面では音楽プレイヤーだけでなく、取得した状況の表示、実験に使用するためのラジオボタン等も表示している。実験に使用する機能については 6.1 節で説明する。

5.3. プレイリストの生成

入力された状況に応じてプレイリストを生成する。プレイリストは 10 曲からなる。最初に各 4 種類の状況に適した楽曲を 15 曲ずつ、合計で延べ 60 曲を選択する。ここからランダムで 10 曲を選択してプレイリストを生成す

る。これは各状況下において常に同じ楽曲が再生されることを防ぐためである。表 2 に、現在地が小金井キャンパス、天気が晴れ、行動が作業中、感情が楽であった場合に推薦されうるプレイリストの例を示す。



図 3 (a)行動選択画面, (b)音楽プレイヤー画面

表 2 現在地が小金井キャンパス、天気が晴れ、行動が作業中、感情が楽であった場合に推薦されうるプレイリストの例

曲順	曲名	アーティスト
1	パプリカ	米津玄師
2	キミに 100 パーセント	きゃりーぱみゅぱみゅ
3	First Love	宇多田ヒカル
4	ふっかつのじゅもん	sumika
5	LOSER	米津玄師
6	napori	Vaundy
7	かわE	ヤバイ T シャツ屋さん
8	ヨワネハキ	MAISONdes
9	ヴァンパイア	DECO*27
10	ただ君に晴れ	ヨルシカ

6. 実験

6.1. 手順

できるだけ異なる状況下で適切にプレイリスト生成が行われ、再生されるかを検証するため、実際に本システムを使用して実験を行った。4.2 節で述べた通り、現在地について、法政大学小金井キャンパス周辺施設の情報だけを取得している。このため、本実験は大学近辺で行った。また、すべての状況下で実験を行うためには 256 通りの状況で実験を行う必要があり困難であると判断したため、以下の 2 通りの方法で実験を行った。

- ① 状況をこちらから指定し、実際にその状況であったと仮定した場合に適切な楽曲がプレイリストに反映されているかを評価してもらう。この際の状況はあくまで仮定であるため、場所の移動や行動は伴わない。作成するプレイリストは 2 曲とする。
- ② 対象者に場所の移動や行動を伴った実験を現地で行ってもらう。この際、感情、天気、行動は固定し、4 か所の場所へ移動することで 4 通りの状況下で本システムを使用してもらう。作成するプレイ

リストには 10 曲が含まれ、それぞれの楽曲が状況に適しているかシステム上で評価する。

本実験の被験者は 14～22 歳の男女 8 人であった。この中に楽曲の特微量に関するアンケートに回答した者は含まれない。①の実験は 1 人当たり 32 通りの状況について評価してもらうことで、8 人全体で 256 通りすべての状況の評価する。②の実験については 1 人当たり 4 通りの条件で実験してもらうため、計 32 通りの結果を得る。

評価は上部に表示されているラジオボタンで、5 段階で行う。適している場合が 5、適していない場合を 1 とする。データ送信ボタンによりシステム内で CSV ファイルを作成して保存する。

6.2. 実験①の結果

256 通りすべての平均スコアは 3.27 ポイントであった。このうち、最高評価である 5 ポイントを取得した状況が 4 通りあり、晴れ／東小金井駅／移動中／楽、雨／小金井公園／運動中／哀、雨／法政大学小金井キャンパス／作業中／喜、雨／法政大学小金井キャンパス／休憩中／哀であった。行動については 4 つすべての行動で 5 ポイントを取得する状況が存在したことから、ユーザーに適した楽曲を選択できていたといえる。

感情別にまとめた評価を図 4 に示す。各感情の平均スコアは、喜が 3.48 ポイント、怒が 2.66 ポイント、哀が 3.38 ポイント、楽が 3.56 ポイントであった。最低評価の選択肢である 1 ポイントを取得した状況のうちその割合が最も大きかったのは、怒であった。怒は、最高評価の 5 ポイントを取得した割合も最も小さく、楽曲選択の基準が誤っていたと考えられる。また、感情の中で最も平均評価が高かったのは楽であった。

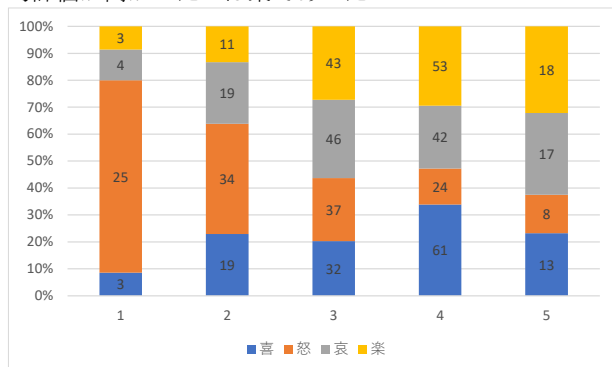


図 4 感情別の評価結果

6.3. 実験②の結果

実験①で得たすべての評価の平均は 3.26 ポイント、実験②で得られた 3.27 ポイントとほとんど差がなかった。被験者ごとの場所別の評価の平均を図 5 に示す。各場所における平均スコアは、法政大学小金井キャンパスで 3.44 ポイント、東小金井駅で 3.28 ポイント、小金井公園で 3.19 ポイント、その他の場所で 3.16 ポイントであった。しかし、東小金井駅での実験で低い評価を得た。この被験者の状況は東小金井駅/晴れ/休憩中/哀であり、この状況の平均スコアが 3.27 ポイントを超えているため、システムに使用した楽曲に問題があったと考える。

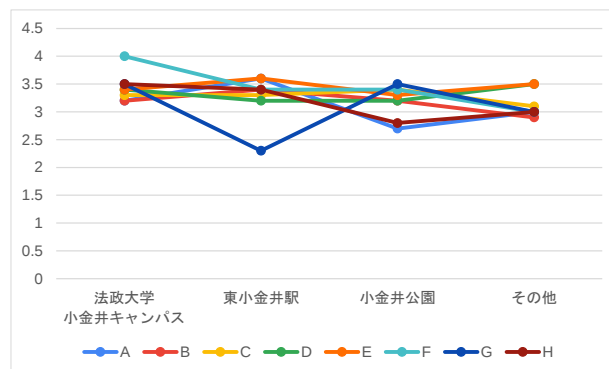


図 5 被験者 A～H による場所別評価の平均

7. 議論

被験者の多くから、楽曲が状況に適しているかどうかは嗜好度に左右されるという意見をもらった。本システムはユーザーの嗜好度までは考慮できておらず、それが課題点になった。今後、状況認識に加え、使用するユーザーの嗜好度も考慮することでより実用的なシステムに改良する必要がある。

知らない楽曲が多いという意見も多かった。システムに使用する楽曲を特微量によって選択したため、有名な楽曲がシステムで使用されない場合があったためであると考えている。また、システムの使用年代を明確に設定しなかったため、本システムに使用する楽曲の年代が様々であった点も原因であると考えている。特微量の数値が強く、どのような状況下でも選択される楽曲が存在した。使用する特微量の種類や数について再検討する必要があると考えている。

8. おわりに

本システムを使用して実験を行った結果、まだ実用性は乏しく、多くの改善点があった。しかし否定的な意見や結果だけではなく、「画期的で楽しめた」「様々な楽曲を知ることができてうれしかった」等の肯定的な意見も多く得られた。実用的なシステムにするために本研究の問題点をさらに洗い出して改善する必要がある。

文 献

- [1] 桐本篤, 佐々木史織, 清木康, "風景画像とサンプル楽曲を用いた環境状況コンテキスト対応型音楽推薦システムの実現," 情報処理学会研究報告: DBS, vol. 146, pp. 157-162, 2008.
- [2] 大出彰男, 安藤彰男, 谷口高士, "音楽によって喚起される感動の種類と音楽や音響に印象の関係," 電子情報通信学会論文誌 A, vol. J97-A, no. 4, pp. 323-331, 2014.
- [3] 加藤内蔵進, 加藤晴子, 逸見学伸, "日本の春の季節進行と季節感を切り口とする気象と音楽との連携: 小学校での授業実践," 天気, vol. 56, no. 4, pp. 203-216, 2009.