

自然な DJ ミックスのための自動選曲手法の提案

大谷研究室

1372026 大野木俊樹

1. 研究の背景と目的

ダンスミュージックとは、ディスコやクラブなどで流れる音楽である。一般的に DJ は 2 つのダンスミュージックのアウトロとイントロの一部またはすべてを重ね合わせて、連続再生する。本研究では、DJ が曲を重ね合わせて連続再生することをミックス、重ね合わせた音が同時に出力される部分をミックス部分、ミックスが施された複数曲が連続的に再生される音源を DJ ミックスと呼ぶ。DJ にとって、ミックスと同等に選曲は大切な要素で、さまざまな要素を考慮しながら選曲する。

現在、楽曲推薦システムが数多く研究されており、音楽アプリケーションなどに応用されている。例として、選択した曲を自動でミックスする Pacemaker が挙げられる。Pacemaker には、現在再生されている楽曲に類似した楽曲を楽曲データベースから抽出し、推薦曲として提示する機能があるが、次曲を自動で決定することはないので、ユーザは提示された候補、あるいは他の曲から手動で次曲を選ばなければならない。また、複数曲が収録されているプレイリストを指定した場合には楽曲のテンポをもとに曲が並び替えられるが、最初の曲が指定できないなど、自由度が低い。さらに、ミックス部分の拍数が定められていないためにミックス部分以外に演奏されるイントロの長さが適切でなかったり、同じ周波数帯の音が同時に出力されることで起こる不協和音によって自然な DJ ミックスが実現できていなかったりすることがある。自然な DJ ミックスを実現するためには、単なる類似曲を次曲として選ぶのではなく、ミックスすることを想定して次曲を選択し、ミックス部分の開始位置と終了位置を適切に決定する必要がある。

本研究では、自然な DJ ミックスのための最適な楽曲の選択を目的として、ミックス部分や楽曲の特徴を考慮した選曲およびミックス部分の決定手法を提案する。

2. 選曲手法の概要

既存セットリスト、五度圏における調間の距離、転調、再生履歴におけるアーティストの情報、歌の有無、およびミックス部分の音の強さに基づいてすべての候補曲の評価値を求め、評価値が最も大きい楽曲を次曲として提示する。すでに選択済みである曲を H_1, H_2, \dots, H_{n-1} として、候補曲 $C_1 \sim C_m$ から H_n を選択するとき、候補曲の C_m 評価値 $Point(C_m)$ を式 (1) により算出する。 w_i は重みで、値を変更することで評価値に与える影響を調整することができる。

$$Point(C_m) = Characteristic(C_m) \times w_0 + Tona(C_m) \times w_1 + Artist(C_m) \times w_2 + Vocal(C_m) \times w_3 \quad (1)$$

$Characteristic(C_m)$ は楽曲に関する特徴の評価値であり、 $Characteristic(C_m)$ は式 (2) で算出される。

$$Characteristic(C_m) = Para_Feature(C_m) \times Para_BPM(C_m) \times Frequency(C_m) \quad (2)$$

$Para_Feature(C_m)$ は楽曲の雰囲気を表す曲想の評価値で、既存セットリストにおいて、曲想が H_{n-1} から C_m と同様に变化する割合である。 $Para_BPM(C_m)$ は楽曲の BPM の評価値で、既存セットリストに

において、BPM の増減値が H_{n-1} から C_m と同様になる割合である。 $Freq(C_m)$ はミックス部分の音の強さである。 楽曲 T_m のイントロとアウトロをそれぞれ $Intro(T_m)$, $Outro(T_m)$, 楽曲 T_m を 32 ビートからなるブロックに分けたときの i 番目のブロックを $Block(T_m, i)$ とし、ブロック b_1, b_2 におけるデシベル増加量を $d(b_1, b_2)$, 実数 x の整数部分を $int(x)$ とする。 $Frequency(C_m)$ は式 (3) で表される。 なお、 C_m が選曲されたとき、 $Frequency(C_m)$ を算出したブロックの組み合わせをミックス部分として提示する。

$$Frequency(C_m) = 1 - \frac{int\left(\min_{j,k}\left(d(Block(Outro(H_{n-1}), j), Block(Intro(C_m), k))\right)\right) \times 100}{301} \quad (3)$$

$Tona(C_m)$ は調に関する評価値である。 楽曲 T_1, T_2 の調の五度圏の円周上における距離を $Dist_Key(T_1, T_2)$ 既存セットリストにおいて、長調または短調が H_{n-1} から C_m と同様に变化する割合を $Para_Transition(C_m)$ とすると、 $Tona(C_m)$ は式 (4) で定義される。

$$Tona(C_m) = \left((6 - Dist_Key(H_{n-1}, C_m)) \div 6 \right) \times Para_Transition(C_m) \quad (4)$$

$Artist(C_m)$ は、同じアーティストの曲が頻繁に選曲されることを避けるために、同じアーティストの選曲数と連続数を候補曲の評価値に反映させる値である。 T のアーティストを $Artist(T)$, $Artist(H_e) \sim Artist(H_f)$ に含まれる $Artist(C_g)$ の個数を $Value(e, f, g)$, $Artist(H_e) \sim Artist(H_f)$ の先頭から $Artist(C_g)$ が連続して現れる個数を $Sequence(e, f, g)$ と表すと、 $Artist(C_m)$ は式 (5) で算出できる。

$$Artist(C_m) = \{10 - Value(e - 11, f - 1, g)\} \times \{10 - Sequence(e - 1, f - 11, g)\} \div 100 \quad (5)$$

$Vocal(C_m)$ は、歌の有無に関する評価値で、既存セットリストの全楽曲において、歌の有無が H_{n-1} から C_m と同様に变化する割合である。

3. 評価実験と考察

DJ2 名とミックス再生に馴染みがあるリスナー3 名に対して、1 曲目をランダムに選択した後、提案手法により 4 回選曲しミックスを施した DJ ミックスを聴かせ、ミックス全体の満足度やミックス部分の感じ方についてのアンケートを実施した。リスナーは満足度およびミックス部分の感じ方について概ね高評価であったが、DJ からは複数の問題点を提示された。DJ からの回答の一部を以下に記す。

- 次々に系統の異なる音楽が流れてくるのでミックスに一貫性がない。
- 急に次の曲が流れてきた。
- 前の曲とテンションが違う。
- ミックス部分を長くしたほうがいい。

本研究では、既存セットリストにおける楽曲の前後関係を基に楽曲に点数をつけるため、ミックス全体の曲の雰囲気推移は考慮しておらず、選曲の一貫性は最初に選ばれた楽曲に依存する。また、すべての候補曲の点数が低いとき、楽曲の特徴以外のパラメータが大きい楽曲が選ばれるので、前の曲と雰囲気が大きく変わることがある。さらに、楽曲の判定基準はすべて点数であるため、どんなに点数が低くても候補曲が 1 つ以上あれば必ず選曲される。そのため、Pacemaker のセットリスト並べ替え機能で並び替えられた最初の 10 曲と、提案手法により選ばれた 10 曲を比較すると、提案手法では BPM が最大で 10 違う楽曲が選択される場合がある。

現状では複数の問題が散見されるものの、本学学生、DJ、およびリスナーの計 16 人に行ったアンケートによると、ミックスを自動で行う機能と自動選曲を組み合わせたシステムが有用であると答えた人は 14 人であった。諸問題を解決することで、有用なシステムが提案できる可能性がある。