

時間指定を考慮したメドレー曲自動生成手法に関する研究 A method of medley generation considering medley's total time

柏田 成[†] 大野 将樹[‡] 獅々堀 正幹[‡]
Akira Kashiwada[†] Masaki Oono[‡] Masami Shishibori[‡]

1. はじめに

本実験の目的は時間を指定した上でメドレー曲を自動生成することである。メドレー曲とは異なる楽曲の断片を連結することによって表現される新たな形式の楽曲のことである。近年では「年代別メドレー」や「アーティスト別メドレー」などジャンル別で作成されたメドレーがカラオケで歌われることが増えたり、ニコニコ動画や YouTube などの動画配信サイトで利用者が自ら作成したメドレーの公開を行ったり、また MIXTRAX[1]といったメドレー楽曲生成システムも普及するなど、社会的にもメドレーという概念が身近に感じられることが増えつつある。

私たちはそれらのメドレー曲は楽曲によって再生時間が大きく異なり統一性がない点に注目し、再生時間の指定を行う研究を行った。メドレー曲の再生時間を指定することによって、定められた時間の動画作成支援などの応用が考えられる。

2. 関連研究

我々はこれまで、楽曲間の繋がりを考慮したメドレー曲の自動生成手法に関する研究を進めてきた[2]。従来手法では、Intro→A メロ→サビ→Ending のように楽曲らしい構成のメドレー曲になることを目指す大局的接続と、周波数重心や BPM 等で抽出したユークリッド距離とクロマベクトルのコサイン類似度を特徴量として音響的に自然に接続することを目指す局所的接続の 2 つの指標を用いて自然なメドレー楽曲生成を実現しており、提案手法もこれに基づいてメドレー曲を生成する。

3. 提案手法

3.1 用語の定義

本研究で用いる用語を以下に定義する。

パート : Intro, A メロ, サビ, Ending 等, 我々が音楽を視聴したときに 1 つのフレーズとして感じることができる楽曲の断片のこととする。

メドレー曲 : 次の条件を満たすものをメドレー曲であると定義する。

- Intro で始まり, Ending で終わる
- メドレーの途中で Intro や Ending を含まない
- 全入力楽曲の歌唱のあるパートを 1 つ使用し構成

3.2 システム概要

システム概要を図 1 に示す。特徴抽出部と接続部は岡村らの手法[2]に基づいたものであり、各々、クロマベクトルなどの特徴量の計算と、使用パートとパートの順番の決定

を行う。提案手法として時間指定部を作成し、再生時間における重みを付けることで違和感が少なくメドレー曲を自動生成する手法を考案した。

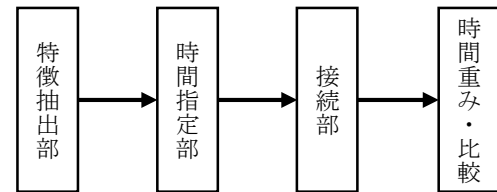


図 1 システム概要

3.3 時間指定部

楽曲の秒数を考慮する手法として最もナイーブな手法は、全曲の各パートの全ての組合せを抽出することだが、楽曲数を N 、楽曲 n におけるパート数を $E(n_i)$ とすると、計算量は $O(\sum_{i=1}^N E(n_i))$ のように累乗的に増えるため現実的ではない。このことを改善するために、提案手法ではパート組合せを最適化問題：ナップサック問題と同等の問題であるとして動的計画法[3]を用いて解決した後、最適なパート順番を決定する。

各パートの演奏時間テーブルを作成した後、パート時間組合せを作成する。前処理として、縦軸がメドレーを構成する楽曲、横軸が再生時間(秒)として表 1 のような時間テーブルを作成する。これにより同じ時間構成のパートを 1 つとして組合せを考えることができ、計算量を小さくできる。

表 1 演奏時間テーブルの例

楽曲	1 秒	2 秒	3 秒	4 秒
楽曲 1		Intro	Chorus	Ending
楽曲 2	Intro		Ending	Chorus

作成した演奏時間テーブルを用いて各楽曲のパート時間組合せテーブルを、以下のような手順で作成する。

- 組合せテーブル 1 段目に楽曲 1 のパートを格納
- for ($i=2; i < \text{楽曲数 } N; i++$)

組合せテーブル $i-1$ 段目と楽曲 i のパートの組合せを組合せテーブル i 段目の適する時間に格納・追加。

指定した秒数を s 秒とすると、計算量はテーブルのセル数 $O(N \times s)$ となり現実的な値となる。表 1 のパート時間構成で 6 秒を指定し組合せを作成した例を表 2 に示す。表の数字と記号は、楽曲 1 の Chorus を 1C のように[楽曲番号][パートの頭文字]を示し、[&]は組合せを表している。

表 2 パート時間組合せテーブルの例

楽曲	2 秒	3 秒	4 秒	5 秒	6 秒
1	1I	1C	1E		
1&2		1I&2I	1C&2I	1I&2E	1I&2C 1C&2E
				1E&2I	

組合せテーブルの最下層 N 段目の指定した時間及び近くのセルに存在する組合せをメドレー曲候補とする。また、

[†] 徳島大学大学院 先端技術科学教育部 Graduate School of Advanced Technology and Science, Tokushima University

[‡] 徳島大学大学院 ソシオサイエンス研究部 Institute of Technology and Science, Tokushima University

組合せテーブル作成の際、全楽曲のイントロが組み合わさるなど、明らかにメドレー曲が完成しない組合せが存在する場合はメドレー曲候補から外すことができる。

3.4 時間重み・比較

最も指定秒数に近いメドレー曲候補を用いてメドレー曲を作成すると、音響的に不自然に繋がるメドレー曲を出力する可能性がある。このことを減らすため、複数のメドレー曲候補に対して再生時間における重みを付けて比較を行うことで、最適なメドレー曲を出力する。本研究では指定時間付近の 5 個程度のメドレー候補を抽出し、各メドレー曲候補について使用パートとパート順番を決定後、各接続におけるクロマベクトルのコサイン類似度の平均などに重みを付けて比較を行い、最終的に出力するメドレーを決定した。重み付けには指定した時間に依存した大きさの窓のハミング窓を用いた。

4. 実験

4.1 実験内容

指定時間に近づいているかという点と従来手法でのメドレー曲よりも違和感が増えているかという点の 2 点について評価実験を行った。楽曲は RWC 研究用音楽データベースのポピュラー音楽 100 曲を使用し、AIST Annotation を用いることによりパート構造を抽出した。

4.2 実験結果

4.2.1 時間に関する評価

楽曲集合を計 80 個作成し、岡村らの手法[2]で作成したメドレー曲 80 曲と提案手法で 2 通り（従来手法±30 秒程度）の時間を指定して用意したメドレー曲 160 曲とで比較を行った。指定した時間と実際に出力されたメドレー曲の再生時間との差異の分布を図 2、どの程度従来手法よりも提案手法の再生時間がどの程度指定時間に近似したかの分布を図 3 に示す。

指定時間との差異の分布を示す図 2 は左にある方が指定時間に近く良い結果であるといえる。98 曲/160 曲のメドレー曲の再生時間が指定時間から前後 5 秒以内となり、良好な結果が出力された。指定時間から離れているメドレー曲が出力された理由としては、指定時間付近にメドレー曲が再生できる組合せが存在しなかったか、存在したが出力したメドレーよりも類似度が低かったという 2 点が挙げられる。

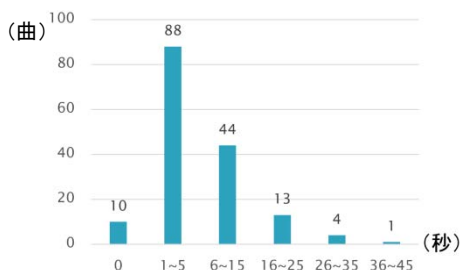


図 2 指定時間と再生時間の差異の分布

近似した秒数の分布を示す図 3 は右にあるほど本手法の効果があったといえる。160 個のメドレー曲のほとんどが指定した時間に 6 秒以上近似したが、図 3 の「~0」の 5 つのメドレー曲では従来手法よりも再生時間が離れてしまっ

た。この原因として時間重みによるものが挙げられる。従来手法でメドレー曲として最適であると判定された組合せも提案手法のメドレー候補として挙がっているが、特徴量の類似度に重みを付けたことより、時間が遠いものが優先された。

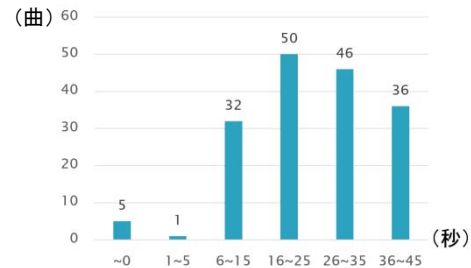


図 3 指定秒数に近似した秒数の分布

4.2.2 違和感に関する評価

ランダムに 10 曲選出した楽曲集合を計 3 個作成し、岡村らの手法[2]で作成したメドレー曲 3 曲と提案手法で作成したメドレー曲を 2 通りの時間を指定して用意した 6 曲で比較を行った。被験者 10 名を選出し 9 曲のメドレー曲について感じる違和感がいくつあるかその個数を判定した。本実験では音楽構成推移か音響的のどちらかを感じる事が出来れば違和感であるとした。結果を各楽曲集合における表 3 に示す。

楽曲集合 A・B・C において違和感の平均個数はそれぞれ減少・増加・ほぼ変化なしという結果が出た。提案手法では従来手法と大きな差が生まれることはなく、接続された楽曲に依存することが考えられる。

表 3 違和感の平均個数

	従来手法	提案手法 A	提案手法 B
楽曲集合 A	3.3	3.1	2.1
楽曲集合 B	2.4	2.7	3.2
楽曲集合 C	3.0	3.0	2.8
手法別平均	2.9	2.817	

5. おわりに

本研究ではパート時間の組合せにおける、時間を考慮したメドレー楽曲生成手法を提案した。今後の課題として以下の点が挙げられる。

- ・違和感であると判断される接続の削減
- ・時間重みの付け方の考察

従来手法でも同様、違和感と判定される接続は一定数存在するため、減らすことが求められる。また、現在クロマベクトルのコサイン類似度などの値に、窓関数における重みを直接掛け合わせることで時間重みとして計算しており、よりよい重み付けの方法を考察するべきである。

指定した時間か違和感が少ないかのどちらかを取るかというトレードインな関係も存在することから、ユーザーによって選択・調整が可能であることが求められる。

参考文献

- [1] Pioneer 社, "MAXTRAX"(2014) <http://www.mixtrax-global.com/jp/>
- [2] 岡村 亮一, 大野 将樹, 沼尾 雅之, "音楽的特徴に基づくメドレー曲の自動生成手法", 第 11 回情報科学技術フォーラム論文集 (2012).
- [3] 草刈 良至, "動的計画法と擬多項式時間アルゴリズム"(2005) <https://www.akita-pu.ac.jp/system/elect/comp1/kusakari/japanese/teaching/InfoMath/2005/note/11/>