

複数の音楽からメドレーを自動生成するシステム

佐藤 学 鈴木 健史 新美 怜志 濱川 礼
中京大学 情報科学部 情報科学科

1. はじめに

本論文では、複数の音楽を半自動的に分割・分析・編集・結合することによって、一つのメドレーを生成するシステムについて述べる。

近年動画共有サイト、特にニコニコ動画[1]で、ユーザによって作成されたメドレーが見られる。しかし、投稿者と自分の好みが完全に一致することは稀であり、自分好みのメドレーを探すことは困難である。自分好みのメドレーを探すことができたとしても、手間や時間を必要とする。また、音楽編集ソフトを用いて手動でメドレーを作成する場合にも、手間や時間を必要とし、経験が乏しい人にとってはハードルが高い。

そこで、本研究では、手軽にユーザ好みのメドレーを生成することで、より多くのユーザがメドレーを楽しむことができると考え、本システムの開発を行った。

2. 本システムの特徴

本システムは、複数の音楽ファイルから半自動でユーザ好みのメドレーを生成する。

ユーザにメドレーの盛り上がり方を入力させ、盛り上がりを元に音楽構成を作成。その音楽構成から自然に聞こえるメドレーを生成することで、ユーザ好みのメドレーを生成できると考えた。

そこで、入力音楽を盛り上がりに沿って並び順を決定し、音楽構成を作成する。また、ボーカル区間の抽出と音楽と音楽を滑らかに繋げることにより、自然に聞こえるメドレーの生成を実現した。本システムでは、ステレオ音源のボーカル有りの音楽を想定している。

3. システムの全体構成

本システムは、分割部、分析部、編集部から構成されている[図 1]。

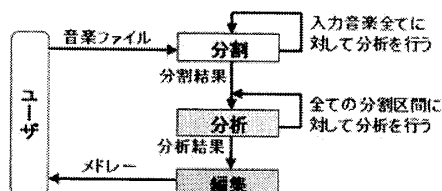


図 1 システム構成図

System that generates medley from two or more music automatically

Manabu Sato, Takeshi Suzuki, Satoshi Niimi
and Rei Hamakawa
Chukyo University Department of information science

4. ユーザインタフェース

[図 2]に本システムのユーザインタフェースを示す。

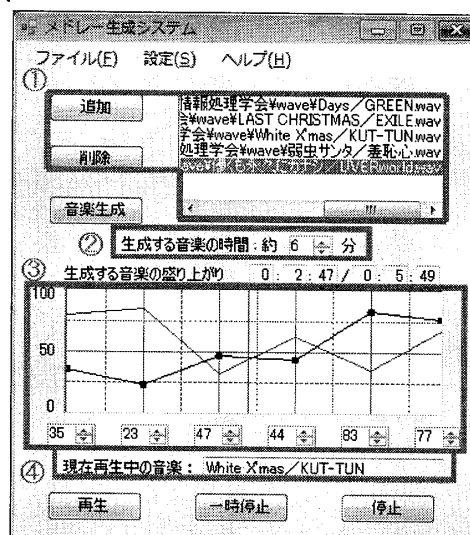


図 2 ユーザインタフェース

①入力音楽ファイルの追加, 削除, 表示

②生成するメドレーの長さ(1~80分)

③盛り上がりのグラフ

黒線: 生成するメドレーの盛り上がり度の入力

青線: 前回の盛り上がりの履歴

赤線: 現在の再生位置

④現在再生中の入力音楽ファイル名

5. 分割部

人の声が途切れることによる違和感をなくするため、ボーカルの有無によって音楽を分割する。

ボーカルの有無を判別するために、ボーカル無しのカラオケデータを作成する。ステレオ音源の場合、ボーカルが中央に、メロディが左右に配置されていることが多いことから、左チャンネルの音と右チャンネルの音の差分を取ることで、カラオケデータを作成する。

元データとカラオケデータを、フーリエ変換を用いて周波数毎に分解し、人の声の周波数帯(80Hz~1000Hz)の音量の総和で比較を行う。元データから閾値を算出し、カラオケデータが閾値を超えた場合、ボーカル有り、超えない場合は、ボーカル無しと判断する[図 3]。

上記の処理を1秒毎に行い、ボーカル無しから有り、また有りから無しに変化した場所で音楽の分割を行う。

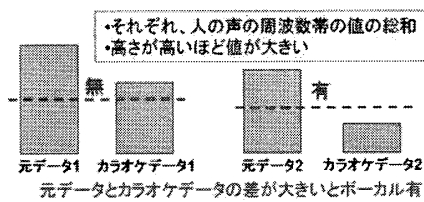


図3 ボーカルの有無の判定

6. 分析部

分割された音楽を分析し、音量・スピード感・音色の3つを分析値として算出する。

音楽が盛り上がっている場合には、音が大きくなる。また、音楽にスピード感があつた。そこで、音量・スピード感を取得し、盛り上がりを再現する。

また、音楽同士で音の高さの違いが大きい場合、音楽と音楽の繋ぎ目の部分で違和感を覚える。そこで、違和感を少なくするために音色を取得する。

6.1. 音量の分析値

分割された音楽の全体の音量の平均を求める。

6.2. スピード感の分析値

著しく音量が高くなっている場所の平均出現間隔を求める。間隔が短いと、1つの音として知覚される。また、間隔が長いと、連続した音として知覚できないため、0.15秒～1.5秒の範囲で算出している[2]。

6.3. 音色の分析値

分割された音楽の始まりと終わりを、フーリエ変換で周波数毎に分解し、各周波数の音量を調べる。分解したデータを低音域・中音域・高音域で区別し、各音域で総和を求め、それぞれを低・中・高音域の音色の分析値として算出する。

7. 編集部

UIで入力した盛り上がりの値と分析部で算出した各分析値を用いて音楽の並び順の決定と、盛り上がりのグラフに沿って音量を変化させる処理を行う。並び替えの簡単な流れを図4に示す。

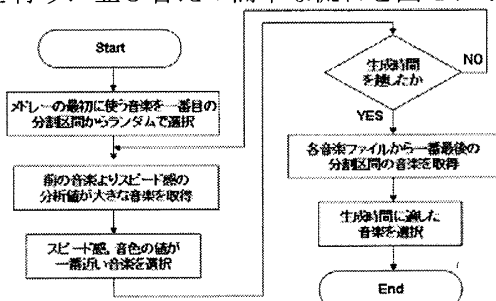


図4 盛り上がりが増える時の並び替えの流れ

メドレーの先頭に配置する音楽は、分割された音楽のうち、元々の音楽で先頭区間にあつた音楽

の中からランダムで決定する。

メドレーの二番目以降に配置する音楽は、盛り上がりが増えるならば、前の音楽よりスピード感のある音楽、下向きならば前の音楽よりスピード感のない音楽の中からスピード感、音色の分析値が一番近い音楽を選択する。また、該当する音楽が無い場合、スピード感の分析値が一番近い音楽を選択する。

メドレーの最後に配置する音楽は、分割された音楽のうち、元々の音楽で最後の区間にあつた音楽の中から二曲目以降と同様の方法で選択する。

そして、並び替え処理後、盛り上がり増えと音量の分析値から音量の変化倍率を計算し、メドレーの音量を増えに沿って変化させる。

8. 評価

本学部の学生35名を対象として、本システムのアンケート形式の調査を行った[図5]。

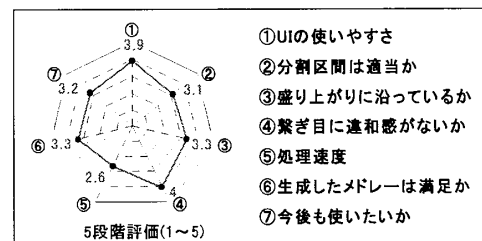


図5 システムの評価結果

「使いやすい」「面白い」「繋ぎが自然」等の意見から、手軽にユーザ好みのメドレーを生成するという目的を達成できたと言える。しかし、5曲で10分のメドレーを生成する場合でも、約3分も時間がかかるということもあり、処理速度については必ずしも良い評価は得られなかった。

9. 今後の課題

今後の課題として、本研究はユーザ好みのメドレーを生成することが最大の目的のため、よりユーザ好みのメドレーを生成する必要があると考える。そこで、メドレーで使用する音楽の優先度を設定する機能と、生成したメドレーの一部のみを変更する機能を追加することで、よりユーザ好みのメドレーの生成が可能であると考えている。

参考文献

- 「ニコニコ動画」
<http://www.nicovideo.jp/>
- 末富大剛, 中島 祥好: ”リズム知覚研究の動向”, 音楽知覚認知研究, Journal of Music Perception and Cognition, 1998, Vol.4, No.1, 26-42