様　式　C－19、F－19－1、Z－19　(共通)

科学研究費助成事業　　研究成果報告書

平成 29 年　 5 月 25 日現在

機関番号：25403
研究種目：基盤研究(C)（一般）
研究期間：2013～2016
課題番号：25330436

研究課題名（和文）音楽における多声性と階層性を表現できるモデル・手法の開発

研究課題名（英文）Models and methods for polyphonical and hierarchical structures in music

研究代表者
林　朗（Hayashi, Akira）

広島市立大学・情報科学研究科・名誉教授

研究者番号：60240909

交付決定額（研究期間全体）：（直接経費）　　1,700,000円

研究成果の概要（和文）：楽曲の階層性を表現できるモデル・手法のベンチマーク問題として，オーディオからの楽曲のセグメンテーションというタスクを取り上げた．オーディオからの楽曲セグメンテーションは，特徴ベクトルの新奇性に基づくアプローチ，特徴ベクトルの同質性に基づくアプローチ，時系列の繰り返しに基づくアプローチの３つに大別できる．従来研究では，これらのアプローチを組み合わせた手法はほとんどない
我々は３アプローチのそれぞれに対応するコスト項目を持つコスト関数に基づく複数アプローチの組み合わせ手法を開発した．実験で，複数のアプローチを組み合わせることにより，セグメンテーションの精度が向上することを示した．

研究成果の概要（英文）：There are three major approaches to music segmentation: novelty-based, homogeneity-based, and repetition-based approaches. However, little work has been done so far to combine all of the three approaches.
We have developed a combined cost function-based approach in which the cost function has cost terms corresponding to each of the three approaches. We show in the experiments using RWC-POP and the Beatles datasets that the accuracy of segmentation is increased by combining the approaches.

研究分野：パターン認識

キーワード：音楽情報検索　楽曲セグメンテーション　コスト関数　２パス・アルゴリズム　ダイナミック・プログラミング

1　研究開始当初の背景

オーディオ・データからの楽曲の構造解析の研究は盛んに行なわれている．オーディオ録音をセグメントに分割し，これらのセグメントを音楽的に意味のあるカテゴリーにグルーピングすることが目標である．従来研究として，ダイナミックプログラミング（DP）を用いた手法がいくつか提案されている．これらの手法では HMM の最尤状態系列の推定に用いられているビタービ・アルゴリズムによく似た DP を用いているが，もとの HMM にセグメントという（状態より上の）階層がないため，セグメント間の類似性の計算など楽曲にある階層性を考慮した構造解析ができないという問題点がある．

ところで，われわれは HMM の状態空間を階層化した階層 HMM については長い研究経験を有しており，最近，階層 HMM の上層の最尤状態系列を求める周辺化ビタービ・アルゴリズムを開発した (1)．このアルゴリズムは階層構造をもつ時系列データを対象とした DP の一種であり，セグメントごとに下層経路確率の周辺化を行うものであるが，同様のアルゴリズムが楽曲構造解析にも適用することにより．従来手法では考慮できなかったセグメント間の類似性を考慮できる DP アルゴリズムが開発できるものと期待される．

2　研究目的

楽曲の階層性を表現できるモデル・手法のベンチマーク問題として，オーディオからの楽曲のセグメンテーションというタスクを取り上げて，オーディイオからの楽曲セグメンテーションのための，コスト関数に基づく複数アプローチの組み合わせ手法を開発する．

サーベイ論文 (2) によると、オーディオからの楽曲セグメンテーションは，特徴ベクトルの新奇性に基づくアプローチ，特徴ベクトルの同質性に基ヲくアプローチ時系列の繰り返しに基づくアプローチの 3 アプローチに大別できる．従来の手法では，これらのアプローチ（の長所）を組み合わせた手法はほとんどない. 唯一の例外として，確率的適合関数によるアプローチがあるが，（Ｉ）最適セグメンテーションは．探潔アルゴリズムによって求められるので， 計算負荷が大きい（２）特徴ベクトルの野崎性に基づくセグメント境界候補の選択が全体の解に大きく影響する．などの問題点がある．

これに対し，我々は， 3 つのアプローチに対応した項の重み付け合計からなる複雑なコスト関数に関して，ダイナミック・プログラミングを用いて（コストの低い）上位 N 候補を効率的に選ぶパス 1 と，上位 N 候補を正薩に比較し，最もコストの低いものを選ぶパス 2 から構成される 2 パスアルゴリズムを開発し，従来手法のような問題点を持たない手法を提案する．

3　研究の方法

3.1　コスト関数

セグメントのコスト関数 $C(S)$ を以下のように定義する。

$$C(S) = \sum_{k=1}^{K} c(s_k) \tag{1}$$

$$\begin{aligned} c(s_k) = {} & w_{nov} \cdot c_{nov}(s_k) + w_{hom} \cdot c_{hom}(s_k) \\ & + w_{rep} \cdot c_{rep}(s_k) + w_{fix} \cdot c_{fix}(s_k) \end{aligned} \tag{2}$$

$C(S)$ は楽曲の総コスト、K はセグメント数、s_k は k 番目のセグメント、$c(s_k)(1 \le k \le K)$ はセグメント s_k のコスト、$c_{nov}(s_k)$,$c_{hom}(s_k)$,$c_{rep}(s_k)$ は、それぞれ新奇性、同質性、繰り返しに基づくコスト項目、$c_{fix}(s_k)$ は固定されたコスト項目、$w_{nov}, w_{hom}, w_{rep}, w_{fix}$ は重みパラメータである。総コスト $C(S)$ を最小化するセグメンテーションを、ダイナミックプログラミングによって求める。

$c_{nov}(s_k)$ は、以下の手順で定義する。

(1) コスト項目の特徴ベクトルを正規化する。$\{\mathbf{f}_{nov}(t) \mid 1 \le t \le T\}$
(2) 正規化された特徴ベクトル間のコサイン距離を用いて、類似度行列を計算する。
(3) 類似度行列の主対角線に沿って、類似度行列の部分行列と、ガウシアンテーパを用いたサイズ k の市松模様のカーネル行列との相関を計算して、新奇性曲線を得る。
(4) 新奇性曲線の値を正規化して、$Nov(t)(Nov(t),0 \le Nov(t) \le 1)$ を求める。
(5) 少なくとも h の振幅を持ち、d 秒離れた、突出した頂点を見つけ出す。
(6) $t = t^end$ のとき $Nov(t)$ が突出した頂点を持つ

ならば、$c_{nov}(s_k)$=0 とする。
それ以外ならば、$c_{nov}(s_k)$=1 とする。

$c_{hom}(s_k)$ は、以下のように定義する。

$$c_{hom}(s_k) = \sum_{t=t_k^{start}}^{t_k^{end}} \|\mathbf{f}_{hom}(t) - \overline{\mathbf{f}}_{hom}^{k}\|^2 \quad (3)$$

$$\overline{\mathbf{f}}_{hom}^{k} = mean(\mathbf{f}_{hom}(t_k^{start} : t_k^{end})) \quad (4)$$

・$\mathbf{f}_{hom}(t)$：時刻 (フレーム)t における同質性に基づくコスト項目の特徴ベクトル

$c_{rep}(s_k)$ は、以下のように定義する。

$$c_{rep}(s_k) = min_{1\le k' \le K, k' \ne k} \{DTW(\mathbf{f}_{rep}(t_k^{start} : t_k^{end}), \mathbf{f}_{rep}(t_{k'}^{start} : t_{k'}^{end}))\}$$

・$\mathbf{f}_{rep}(t)$：時刻 (フレーム)t における繰り返しに基づくコスト項目の特徴ベクトル
・$DTW(\cdot, \cdot)$：daynamic time warping によって計算された、与えられたベクトル系列間の距離

$c_{fix}(s_k)$ は、以下の様に定義する。

$$c_{fix}(s_k) = 1 \quad (5)$$

3.2　２パスアルゴリズム

パス１ではコストの近似計算を行い、セグメンテーション解の候補を求める。セグメントの類似度にも基づくコストは、与えられたセグメントと、重ならない同じ長さの部分時系列との二乗距離の最小値をとり、総コストの近似計算を行っている。

パス２では、パス１で求めた解候補について、正確なコストを計算し、最小のコストを持つものを推定結果とする。セグメントの類似度に基づくコストは、与えられたセグメントと他のセグメントとの DTW 距離の最小値とする。

4　研究成果

4.1　実験

３つの実験を行った．それぞれの実験の目的は以下のとおりである．

- 実験１：　コスト関数にコスト項目を付け加えることによりセグメンテーションの精度が向上することを示すことにより，組み合わせアプローチの優位性を証明する．
- 実験２：１パス・アルゴリズムと２パス・アルゴリズムを比較し，２パス・アルゴリズムによりセグメンテーションの精度が向上することを示す．
- 実験３：新奇性，同質性，繰り返しに基づくコスト項目以外のコスト項目を新たに付け加えることにより，セグメンテーションの精度が更に向上することを示し，提案するアプローチのポテンシャルを証明する．

The Beatles の 174 曲と、Tampere University of Technology(TUT) の正解データ (3) を用いて楽曲のセグメンテーションを行った。

実験では、17 曲で訓練し、F 値の平均値が最大になる重みパラメータを探し出し、その重みパラメータの値を用いて、残りの 157 曲のセグメンテーションを行った。セグメント境界の許容誤差は 3 秒とした．

図１に実験１の結果を示す．図中の nov, hom, rep はそれぞれ新奇性，同質性，繰り返しに基づくコスト項目を表す．図中の数字は 157 曲の F 値の平均値である．コスト項目を付け加えることにより，F 値の平均値が向上していることがわかる．

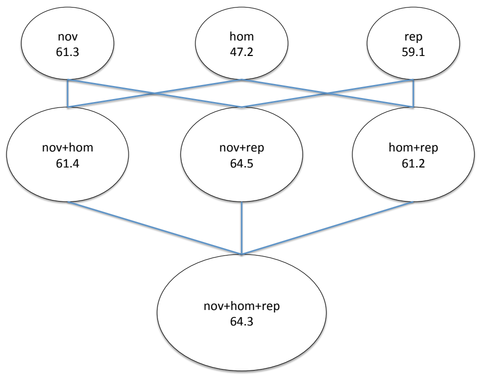

図１　Average F(%) values for the two-pass algorithm (Beatles)

表１に実験２の結果を示す．４種類のコスト関数に対して，いずれも１パス・アルゴリズムより２パス・アルゴリズムの方が，より高い精度（F 値）であることがわかる．

表２に実験３の結果を示す．セグメント長さに基づくコスト項目を付け加えることにより，更に精度が向上していることがわかる．

＜引用文献＞

(1) Akira Hayashi, Kazunori Iwata, Nobuo Suematsu, ”Marginalized Viterbi algorithm for

コスト関数	1 pass	2 pass
$\tilde{c}3$ *vs* $c3$ (rep)	58.0	**59.1**
$\tilde{c}5$ *vs* $c5$ $(nov + rep)$	60.7	**64.5**
$\tilde{c}6$ *vs* $c6$ $(hom + rep)$	59.9	**61.2**
$\tilde{c}7$ *vs* $c7$ $(nov + hom + rep)$	61.2	**64.3**

表 1 Average F(%) values for One-Pass and Two-Pass algorithms

コスト関数	P(%)	R(%)	F(%)
nov+hom+rep	68.8	62.2	64.3
nov+hom+rep+len	71.4	64.3	**66.8**

表 2 Adding the length-based cost term N=5000

hierarchical hidden Markov models", Pattern Recognition, 46, pp.3452-3459, 2013.
(2) J.Paulus et al., "State of the Art Report: Audio-Based Music Structure Analysis", Proc.ISMIR,2010.
(3) Structure annotations for Beatles songs. `http://www.cs.tut.fi/sgn/arg/paulus/beatles\_sections\_TUT\.zip`

5 主な発表論文等

- 〔学会発表〕（計 1 件）
  森本康徳，林朗，末松伸朗，岩田一貴，2 パスアルゴリズムによる楽曲のセグメンテーション，電子情報通信学会綜合大会，2015 年 3 月，立命館大学びわこ・くさつキャンパス（滋賀県草津市）
- 〔その他〕
  古川健太郎，コスト関数に基づいた組み合わせアプローチによる楽曲のセグメンテーション，卒業論文，広島市立大学情報科学部知能工学科，2016

6 研究組織

1. 研究代表者: 林　朗　(Hayashi Akira)　広島市立大学・情報科学部・名誉教授　研究者番号：60240909