

口ドラム認識手法とそのドラム譜入力システムへの応用 音楽情報処理の最近の研究

岩淵 勇樹

1 口ドラム認識手法とそのドラム譜 入力システムへの応用

Abstract

本論文では、人がドラムの音を「ドンタンドタン」のように擬音語で真似た音声（口ドラム）を、それに対応するドラムパターンとして認識する手法を提案する。応用例として、口ドラムによるドラム譜入力インタフェース Voice Drummer を実装した。

1.1 はじめに

音声（歌唱）から対応メロディや楽曲を得る研究としては、ハミング（鼻歌）や歌声を用いた音楽検索手法がある [1-3] が、ドラムパートは音色の違いこそが重要な要素であるため、本研究にそのまま適用することは難しい。

1.2 口ドラム認識手法

認識処理の具体的な流れを図1に示す。

1.3 口ドラム表現実験

19歳から31歳の男女17人の被験者による口ドラムの歌唱を収録して分析した。対象とするドラムパターンは、Bass Drum (BD) と Snare Drum (SD) のみで構成される。

1.3.1 実験方法

図2の10パターンのドラムパターンを楽譜も見ながら被験者に聞かせ、記憶するまで何度も聴取する。その後、再生を停止し、被験者は自分で擬音語表現を考えて歌唱する。

1.3.2 実験結果

得られた結果を考慮した最終的な発音辞書を表2に示す。

1.4 口ドラム認識実験

1.4.1 評価方法

パターン評価と発音時刻評価の2種類の評価を行う。パターン評価は、認識の結果得られたシーケンス（図1、A4）とドラムパターン（A6）が、それぞれ正しく推定されたかどうかを評価する。

発音時刻評価は、正解時刻と認識された発音開始時刻（図1、A4）との「ずれ」を評価する。

1.4.2 実験結果

(1) パターン評価 表3にシーケンスおよびドラムパターンの認識率、表4にシーケンスのN-best 認識率を示す。図3に、被験者ごとの認識率のグラフを示す。図4には、ドラムパターンごとの認識率のグラフを示す。

(2) 発音時刻評価 図5に認識された発音開始時刻のずれの頻度分布を示す。

1.5 ドラム譜入力インタフェース：Voice Drummer

口ドラムでドラムパターンを入力することができる新しい楽譜入力インタフェース Voice Drummer を実装した。実装には、Microsoft Visual C++ .NET 2003を使用した。図6にその画面例を示す。

1.6 おわりに

2 音楽情報処理の最近の研究

2.1 音楽情報処理が実世界と結び付く

最近の音楽情報処理は処理対象が個々の楽曲の内部（音符や和音等）から楽曲全体やその集合へと大きくなり、音楽知識がなくとも利用可能な音楽システムも活発に研究されている。

2.2 楽曲の集合を扱う

音楽情報検索 (MIR: Music Information Retrieval) という、楽曲を一単位として大量の楽曲の集合に対する検索、分類、管理に取り組む研究は、国際学会 ISMIR が 2000 年以降毎年開催される等、活発に研究されている。ここでは単なる書誌情報 (CDDB による曲名やアーティスト名) に基づくテキスト検索ではなく、音響信号に基づく検索に焦点を当てて、以下の三つに分類して紹介する。

2.2.1 メロディを扱う研究: ハミング検索 (QBH: Query by Humming)

聞いたことのある曲を「ラララー」等のように口ずさむと、その曲名を検索できる方法である。メロディのみの場合には、検索キーとの類似度を直接求めればよいが、SMF (標準 MIDI ファイル) の場合には、どのトラックがメロディかを推定してから、類似度を求めなければならない。更に音響信号の場合には、混合音中のメロディとの類似度を求める必要があり、より実現が難しい。

2.2.2 楽曲の断片を扱う研究: 断片を含む楽曲の検索

町中で流れている音楽の曲名を知りたいときに、携帯電話でその一部を録音すると、曲名を検索できる方法である。ベクトル量子化されたパワースペクトルの形状のヒストグラムに基づく時系列アクティブ検索法や、パワースペクトルのピークの出現パターンに基づく方法などが提案されている。

2.2.3 楽曲全体を扱う研究: 楽曲間の類似度に基づく検索

ある曲に似た曲調の曲を探し出すことができる方法である。楽曲間の類似度を、楽曲中の音色 (パワースペクトル形状)、リズム、変調スペクトル、歌声等の様々な特徴に基づいて定義する必要がある。

2.3 音楽音響信号を理解する

音楽の音響信号を理解する研究は過去 10 年間に大きく展開した。

各種統計的手法は広く導入され、隠れマルコフモデル (HMM) 等の確率モデルに基づく手法や、最尤推定、ベイズ推定等を活用した様々な手法が提案されている。

2.3.1 音源分離・音高推定

2.3.2 音楽情景記述

メロディ、ベース

ビート

サビ、フレーズの繰り返し、楽曲構造

楽器の音色 混合音中の個々の音の楽器名を、音高と同時に推定する音源同定手法や、ドラム音の発音時刻を推定する手法が実現されるようになった。

2.4 楽曲のメタ情報を扱う

2.4.1 利用

2.4.2 抽出

2.4.3 記述と標準化

2.5 エンドユーザが使うことを考える

2.5.1 実世界指向

- musicBottles: ガラス瓶を各トラックの再生に割り当てる
- FieldMouse: ID タグを空中で動かして選曲やボリューム変更の操作をする
- SmartMusicKIOSK: 曲のサビまでスキップできる

2.5.2 擬音語

「ストトン音楽」表記システムでは、「ドレミーソツミレド」のようにテキストで旋律を記述するとそれが演奏される。口ドラム検索手法では、「ドンタンドドタン」のように口ずさんだドラムパートを楽曲検索に用いる。

2.5.3 コミュニケーションツール

- CosTune: ウェアラブル楽器
- Music Resonator: 楽曲断片を加工して共有
- RemoteGIG: インターネットによる遅延を配慮し、1周期分遅れたデータと共に合奏する

2.6 音楽情報処理は今後更に発展する

出典

- [1] 中野 倫靖, 緒方 淳, 後藤 真孝, 平賀 譲: 口ドラム認識手法とそのドラム譜入力システムへの応用, 情報処理学会論文誌 Vol. 48 No. 1, pp. 386-397, 2007
- [2] 後藤 真孝, 平田 圭二: 音楽情報処理の最近の研究, 日本音響学会誌 Vol. 60, No. 11, pp. 675-681, 2004