

ドラムパターンの大域的構造と 音響特徴量パラメタの相互作用

澤田 隼^{1,a)} 深山 覚² 後藤 真孝² 平田 圭二¹

概要：本稿では、記号処理による大域的なドラムパターン構造の認識結果を、信号処理によるドラムパターンの認識にフィードバックする枠組みを提案する。従来のドラムパターンの差し替えシステムでは、小節単位の信号処理によるドラムパターンの認識と、記号処理による大域的なドラムパターン構造の認識を別々に行っていた。大域的なドラムパターン構造が既知のときには、小節単位のドラムパターンの認識精度を改善できる。しかし記号処理による大域的なドラムパターン構造の認識結果を用いて、信号処理による小節単位のドラムパターン認識の精度を向上できる統合的な枠組みがなかった。そこで本稿では、一般的にドラムパターンが有する性質に基づいた目的関数を設計し、その目的関数の値を最大化するように、記号処理によるドラムパターン構造の認識結果を信号処理による小節単位のドラムパターン認識へとフィードバックすることで、ドラムパターンの認識性能を改善する枠組みを提案する。本枠組みを検証する実験を行い、実際にドラムパターンの認識性能を向上できることを確認した。

1. はじめに

音楽情報処理の分野では、楽曲の構造を分析する研究が多数存在する [8], [9], [15]。楽曲構造の分析は、適切な楽曲の編曲 [10], [11], [16] や楽曲の表情付け [7], [12], [13], [14] といった応用タスクには不可欠である。近年では楽曲の自動生成に関する研究が盛んに行われているが、高水準の楽曲を生成するためには、何らかの楽曲の構造に関する情報を学習、あるいは組み込む必要がある。

従来の我々の試みでは、ドラムパターンの遷移確率を用いて大域的なドラムパターンをモデル化した [17]。またそのモデル化に基づいて、大域的な構造を考慮したドラムパターンの差し替えシステムを実現した。このシステムには、調波・非調波音分離 (Harmonic/Percussive sound separation: HPSS) [3], [4] を援用した小節単位のドラムパターンの認識と、大域的なドラムパターンの認識という2つのパターン認識プロセスが存在する。大域的なドラムパターンの認識の精度は小節単位のドラムパターンの認識の精度に依存する。しかし、大域的なドラムパターンの認識結果を用いて、小節単位のドラムパターンの認識精度を向上させる枠組みがなかった。

小節単位のドラムパターンは、それを認識する際に用いる特徴量の抽出器のパラメタによって認識精度が変化する。伴奏を構成する楽器や演奏スタイルの違いにより、楽曲を構成する音の特性が変化することで、調波・非調波音分離に最適なパラメタが異なるため、楽曲ごとにパラメタを調節する必要がある。例えば、伴奏とドラムの周波数帯域が近接しドラムがスペクトル中で非顕在である場合は打楽器などの非調波成分とその他の調波成分を分離する閾値を変更する必要がある。

人間が音楽を理解する際、楽曲を聴取し、経験等を通じて新しい知識を獲得すると、それに基づいて楽曲の認識を変化させ、さらに高度な認識や知識の獲得ができるようになるという相互作用が働いていると考えられる。計算機システムに音楽を人間のように理解させるためには、このような信号レベルの認識と、記号レベルの推論の相互作用フレームワークが必要であると考えられる。本研究では音楽理解のための記号処理と信号処理の相互作用フレームワークの構築を目指す。本稿では、記号処理の知識に基づく目的関数を定義し、信号処理にフィードバックすることでHPSSのパラメタを最適化し、小節単位のドラムパターンの認識性能の向上を目指す (図 1)。

本稿では、初めに従来の大域的なドラムパターンのモデル化の方法について述べ (2 節)、続いて大域的なドラムパターンが有すると考えられるドラムパターン間の遷移に関する目的関数を定義し、大域的なドラムパターン構造の認

¹ 公立はこだて未来大学
Future University Hakodate

² 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology (AIST)

a) b1012046@gmail.com

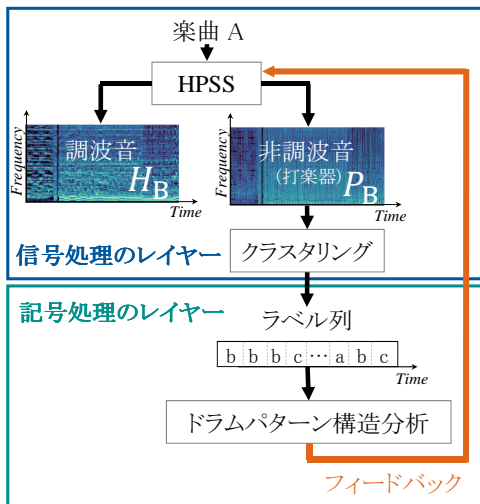


図 1 研究アプローチ

識結果を、信号処理による小節単位のドラムパターンの認識処理へとフィードバックする方法を議論する(3節)。また、3節で定義した目的関数を用いて HPSS のパラメタを最適化し、ドラムパターンのクラスタリングの精度が向上することを確認することで、目的関数の有用性を検証する(4節)。

2. 大域的なドラムパターンのモデル化

典型的なドラムトラックは、複数のドラムパターンが連結されることで形成される大域的なパターン構造を有している。例えば、A メロや B メロ、サビなどの楽曲構造によってドラムパターンが変化する。また、楽曲の構造の変わり目には、今まで繰り返していたドラムパターンとは異なるドラムパターン(フィルイン)が挿入される。ドラムパターン間の遷移確率に着目してみると、楽曲内で頻繁に繰り返される基本パターンは自己遷移確率が高い。またフィルインは特定の基本ドラムパターンから遷移し、自己遷移することなく次のパターンへ遷移するため、特定のパターン間の遷移確率が高くなる。従って、そのパターンの遷移確率には楽曲内のコンテキストの役割が反映された特徴があらわれると考えられる。この仮説に基づいて、ドラムパターン間の遷移確率分布を用いた大域的なドラムパターン構造のモデル化を行った[17]。

RWC 研究用音楽データベース [5] 中の実際の楽曲(RWC-MDB-P-2001 No.3) のドラムパターン間の遷移確率を図 2 に示す。図 2 中のドラムパターン a, b, c は楽曲内で繰り返される基本パターンであり、実際に自己遷移確率が高いのが確認できる(それぞれ順に 0.86, 0.72, 0.67)。また、図 2 中のドラムパターン d, e, f はフィルインのパターンであり、特定の基本パターンからの遷移確率が高いことが確認できる。例えば、フィルインパターン d は基本パターン b からのみ遷移し、必ず基本パターン b に戻る。基本パター

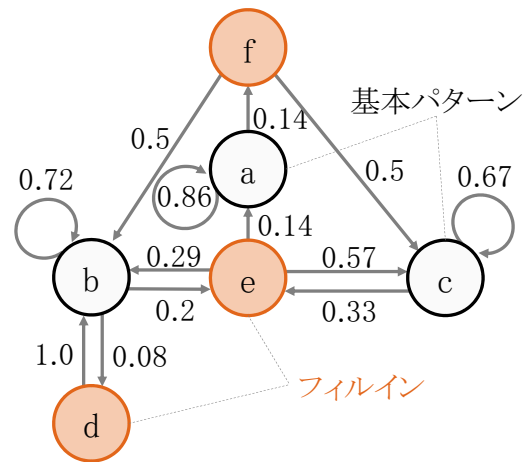


図 2 RWC-MDB-P-2001 No.3 のドラムパターン遷移確率モデル：各ノードの記号がパターンのラベルを表し、ノード同士を結ぶエッジが遷移を表す。矢印の向きにパターンが遷移することを表し、エッジ上に書かれた数字がその遷移確率を表す。ラベル a, b, c が基本パターンのラベルであり、ラベル d, e, f (橙) がフィルインのラベルである。

ン b はフィルインパターン e にも遷移するが、フィルインパターン e はその他の基本パターン a, c にも遷移する。つまり、フィルインパターン d は基本パターン b が繰り返される箇所のみ挿入され、フィルインパターン e は基本パターンが別の基本パターンに変化する箇所に挿入されやすいといった特徴が確認できる。

ドラムパターン間の遷移確率を求めるには、まず一小節単位の非調波音の音響特徴量を求め、その特徴空間でクラスタリングを行い、一小節単位の非調波音(打楽器音)のクラスターを求める。本研究では、ここで得られた各クラスターをドラムパターンと呼ぶ。次に楽曲中の各小節の音響特徴量をドラムパターンのラベルに変換し、このラベルのバイグラム頻度からパターンの遷移確率が求められる。

小節単位のパターンの特徴量を求める際には、まず入力された楽曲の音響信号をスペクトログラムに変換し、ビートトラッキングを行い、一小節毎にスペクトログラムを分割する。次に、一小節毎のスペクトログラムに対してそれぞれ調波・非調波音分離を行い和音やメロディーなどの調波音と打楽器などの非調波音成分に分離する。分離された非調波音成分から、スペクトルの時間的変化を表す尺度として用いられる 1 小節単位のスペクトルフラックス(Spectral flux) [2] の時系列を求める。こうして得られたスペクトルフラックスに主成分分析(Principle component analysis: PCA) [6] を適用し、4次元へと次元圧縮して特徴量として用いる。

3. 音響特徴パラメタへのフィードバック

音響特徴量を抽出するアルゴリズムの多くには、事前に設定する必要があるパラメタが存在する。一般的にそれら

のパラメタは人手で設定される。パラメタの値によって抽出される特徴量の値は変化し、そのパラメタを調節することで分析対象にアルゴリズムを適応させることが可能になる。一方で、適切なパラメタを設定するためにはその分析対象の知識や観察が必要である。教師あり学習の問題設定では、あらかじめ正解データを用いてパラメタの調整をすることが可能である [1]。しかし、教師なし学習やクラスタリングの場合、その方法を用いることができない。本研究においても、楽曲毎に最適なパラメタが異なる場合と考えられ、パラメタの調整により、より正確にドラムパターンをモデル化できると期待されるが、楽曲毎に人手でパラメタを調節するのは現実的ではない。そこで、大域的なドラムパターンの遷移に関する知識を目的関数として表現し、その目的関数を最大化することで適切なパラメタを獲得する。

本研究では「あるドラムパターンから遷移する先のドラムパターンの候補は少ない」というドラムパターン遷移についての知識を用いることでパラメタの最適化を行う。このことは次に挙げる点から見て妥当であると考えられる。まず、楽曲の中で頻繁に繰り返される基本パターンは自己遷移が多い一方で、フィルインパターンは特定のドラムパターンからのみ遷移し、その遷移頻度は他のパターンの遷移に比べて少ない。また、フィルインパターン自身は自己遷移することが少なく、すぐに他のパターンへ遷移する。実際図 2 のように、楽曲内に存在するドラムパターン間の遷移は全結合ではなく、特定のパターン間の遷移に限られている。

このようなドラムパターン間の遷移に関する知識を導入するため、目的関数を以下のように設計する。

$$\sum_i^N \sum_j^N -\log(P(c_i|c_j) + 1) \quad (1)$$

ここで、 N はドラムパターンの総数、 $P(c_i|c_j)$ はドラムパターン c_j から c_i への遷移確率を表す。また対数の底は 2 である。図 3 からわかるように、 $-\log(P(c_i|c_j) + 1)$ の値は $P(c_i|c_j) = 0$ の時に最大値 0 をとる ($0 \leq P(c_i|c_j) \leq 1$)。したがって、条件付き確率が 0 に近い値を持つ遷移先のドラムパターンが多く含まれ、特定のドラムパターンへの遷移だけが観測される場合に、この目的関数は大きな値を持つ。この目的関数を最大化するように、音響特徴パラメタを変化させると、各ドラムパターンの遷移先の種類が少なくなるようにクラスタリング結果が変化する。

また、設計した目的関数を最大化するように HPSS のパラメタを最適化することで、大域的な構造を考慮したクラスタリング結果の修正も期待できる。一般的に楽曲は、1 サビ、2 サビといった部分構造を有する。構造内のドラムパターン間の遷移は、他の部分構造内のドラムパターン間の遷移とともに、どちらも遷移先が限られた遷移確率を

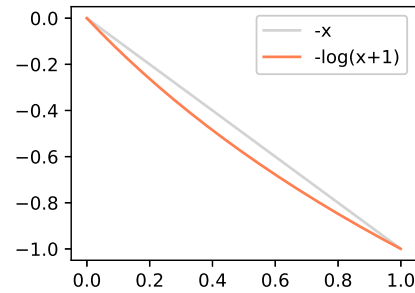


図 3 $-x$ と $-\log(x+1)$

持っていることが多い。したがって、HPSS のパラメタの影響でドラムパターンのクラスタリングの一部を誤ってしまっても、遷移先の種類が少なくなるようにクラスタリングの際の識別面が変化するため、部分構造同士で同じような遷移になるようにクラスタリング結果が修正されることも期待できる。

4. 最適化された音響特徴パラメタを用いたクラスタリングの評価

目的関数の有用性を検証するために、HPSS のパラメタを変化させ、目的関数の最大値をとるときにクラスタリングの精度が高くなることを確認する。本研究で用いた HPSS [3] では、文献 [4] で述べられた手法が拡張され、調波音成分と非調波音成分に分離するマスクの閾値のパラメタ β が導入されている。このパラメタ β が変化すると、HPSS によって得られる分離信号の調波音成分と非調波音成分の性質が変化する。一般的に打楽器は非調波音成分として分離されるが、打楽器の種類や音色によっては非調波音成分と調波音成分の両方を含む。楽曲によって使用されるドラムの音色や、伴奏のスタイル（使用される楽器等）が違うため、楽曲によって適切なパラメタ β は異なると考えられる。本論文ではこの閾値のパラメタ β を式 (1) の目的関数を最大化するように最適化した。

4.1 実験条件

提案手法との比較として、表 1 に示す尺度を用いたドラムパターンのクラスタリングを行った。AllMean はパラメタ β の定義域 ($1.0 \leq \beta \leq 10.0$) 内でパラメタ β の条件を網羅的に変更しながらクラスタリングを行い、パラメタ β 毎に算出される全 F 値の平均値を求めたものである。その楽曲内でパラメタ β を調節して得られる F 値の平均値である。FixParam は全ての楽曲に対して同じパラメタを用いたもので、本論文では HPSS のパラメタとして $\beta = 5.0$ を用いた。CentroidDist は各クラスタのセントロイド同士の距離の和を目的関数としたもので、ドラムパターンの遷移の情報を用いないクラスタリングのクラスタ間の違いを表した尺度である。AllMean と FixParam はパラメタの

表 1 実験に用いる目的関数一覧

β の最適化	目的関数名	説明
なし	AllMean	全パラメタの F 値の平均値
	FixParam	固定パラメタ
あり	CentroidDist	クラスタのセントロイド間の距離
	Entropy	ドラムパターン間の遷移確率のエントロピー
	SparseCount	パターン間の遷移確率が 0 のエッジの数
	Proposed	目的関数 (式 1)

最適化を行わない場合、CentroidDist はクラスタリングの結果のみを用いてパラメタを最適化し、遷移確率による大域的な構造を考慮しない場合に対応する。Entropy はドラムパターン間の遷移確率のエントロピーであり (ただし符号が逆符号)、クラスタ c_j からクラスタ c_i への遷移確率 $P(c_i|c_j)$ を用いて次の式で計算される。

$$\sum_i^N \sum_j^N P(c_i|c_j) \log P(c_i|c_j) \quad (2)$$

この値は、ドラムパターンの遷移先候補が複数に散らばっている場合には低くなり、遷移確率 $P(c_i|c_j)$ が多くの i について 0 である時に大きくなる。Entropy と Proposed を比較し、それらの特性の違いを調べる。SparseCount は、ドラムパターンの遷移確率分布のうち遷移が観測されない (遷移確率が 0) ノードのペアを数え上げ、その個数を目的関数の値として用いた。全てのドラムパターンが特定の 1 つのドラムパターンにしか遷移しない場合に最大値 ($N \times N - N$) をとる。

ドラムパターンのラベルは教師なしクラスタリングの結果として得られるものであり、ラベルとドラムパターンの対応付けには任意性がある。したがって、手法によって自動的に付与されるラベルと正解データで与えられているラベルは必ずしも一致しない。そこで F 値を計算するために必要な再現率と適合率を次のように計算した。まず、手法によって得られるラベルの総数と、1 曲中で正解データとして付与されたラベルの総数が同じである場合のみ、F 値を計算するようにした。この時、再現率は正解データのあるラベルが付与されたドラムパターンのうちで、手法によって出力されたラベルの数を比べた際に、最も数の多いラベルが付与されたドラムパターンが占める割合とした。同様に、適合率は手法によって出力されたあるラベルが付与されたドラムパターンのうちで、正解データのラベルの数を比べた際に、最も数が多いラベルが付与されたドラムパターンが占める割合とした。この再現率と適合率の調和平均をとったものを F 値とする。

本実験では、RWC 研究用音楽データベース (RWC-MDB-P-2001) [5] の楽曲を使用した。そのうち、クラスタ数が 6 つの楽曲 (RWC-MDB-P-2001 No.5, No.7, No.8, No.10, No.12, No.13, No.14, No.15, No.17, No.18, No.24) 計 11 曲を用いた。ドラム演奏経験が 3 年以上ある 2 名がクロス

表 2 目的関数の F 値

目的関数名	再現率	適合率	F 値
AllMean	0.45	0.45	0.45
FixParam	0.45	0.45	0.45
CentroidDist	0.50	0.49	0.49
Entropy	0.53	0.52	0.52
SparseCount	0.52	0.52	0.52
Proposed	0.54	0.54	0.54

表 3 楽曲毎のそれぞれの目的関数の F 値 (Top 3)

目的関数名	楽曲番号		
	No.7	No.8	No.5
AllMean	0.50	0.55	0.46
FixParam	0.52	0.54	0.43
CentroidDist	0.50	0.67	0.52
Entropy	0.73	0.66	0.52
SparseCount	0.73	0.66	0.52
Proposed	0.73	0.66	0.63

表 4 楽曲毎のそれぞれの目的関数の F 値 (Worst 3)

目的関数名	楽曲番号		
	No.15	No.17	No.13
AllMean	0.43	0.38	0.39
FixParam	0.34	0.44	0.38
CentroidDist	0.54	0.44	0.43
Entropy	0.34	0.38	0.45
SparseCount	0.54	0.38	0.45
Proposed	0.34	0.38	0.45

チェックして作成したラベル列を正解データとして用いた。

4.2 実験結果

表 2 に実験に使用した楽曲 11 曲の F 値の平均値を示す。AllMean と FixParam の F 値は 0.45 であり、CentroidDist の F 値が 0.49 であるのに対して、Proposed の F 値が 0.54 となり、パラメタを最適化しない場合 (AllMean, FixParam) およびパターン間遷移を考慮しない場合 (CentroidDist) を上回り、最大となった。次いで Entropy と SparseCount の F 値が 0.52 となり、パラメタを最適化しない場合 (AllMean, FixParam) およびパターン間遷移を考慮しない場合 (CentroidDist) を上回る結果となった。

次に、楽曲毎の F 値を示す。表 3 に Proposed の F 値が高い 3 例を、表 4 に F 値が低い 3 例の結果を示す。楽曲 RWC-MDB-P-2001 No.7 の F 値は Proposed によって最適化されたパラメタを用いたときの F 値が 0.73 となり、ベースラインの AllMean (F 値 0.50) と FixParam (F 値 0.52)、CentroidDist (F 値 0.50) よりも高い値になった。この時のパラメタの値による F 値と目的関数の変化のグラフを図 4 に示す。F 値が最大になるパラメタ β は 2.5 の時であり、目的関数 Proposed が最大値をとる時に F 値も最大になっている。また、目的関数 Proposed が最小値の時に、F 値も最小になっている (パラメタ $\beta = 1.5$)。

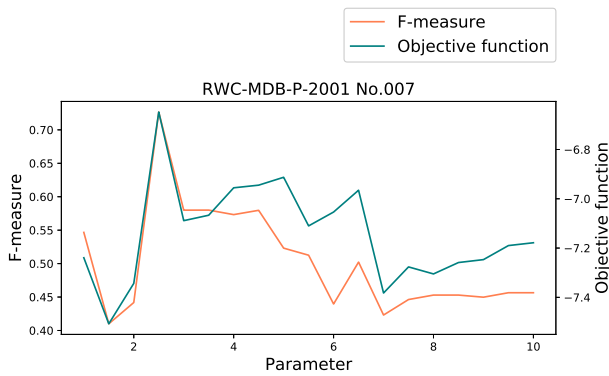


図 4 RWC-MDB-P-2001 No.007 : x 軸が HPSS のパラメタ β , y 軸が目的関数 Proposed の値 (緑) と F 値を表している (橙).

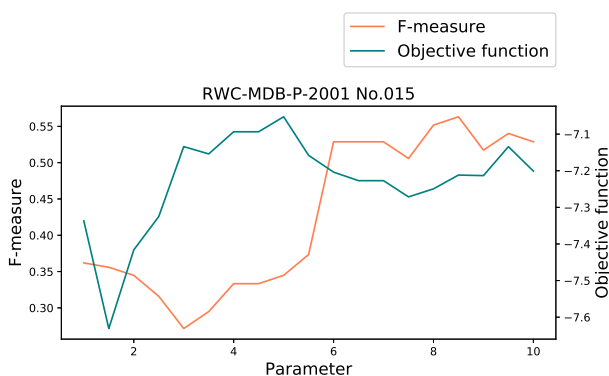


図 5 RWC-MDB-P-2001 No.015 : x 軸が HPSS のパラメタ β , y 軸が目的関数 Proposed の値 (緑) と F 値を表している (橙).

一方で、楽曲 RWC-MDB-P-2001 No.15 の F 値は目的関数 Proposed によって最適化されたパラメタ ($\beta = 5.0$) を用いたときの F 値が 0.34 となり、AllMean (F 値 0.43) や CentroidDist (F 値 0.54) よりも低い値になった。この時のパラメタによる F 値と目的関数 Proposed の変化のグラフを図 5 に示す。F 値が最大になるパラメタ β は 8.5 の時であり、目的関数が最大値をとる時 (パラメタ $\beta = 5.0$) に F 値は最大ではなく、F 値は他の β の条件と比較して小さい値となった。パラメタ β が 1.0 から 6.0 である場合の F 値は比較的低い値であったが、目的関数の値は他の β の条件と比べて高い値であった。

提案法 (Proposed) によって最も高い F 値を実現した楽曲 RWC-MDB-P-2001 No.7 において、F 値と目的関数 Proposed の値が最大であった $\beta = 2.5$ の時のクラスタ間の遷移確率行列を表 5 に、F 値と目的関数 Proposed の値が最小であった $\beta = 1.5$ の時のクラスタ間の遷移確率行列を表 6 に示す。表 5 の遷移確率行列は表 6 の遷移確率行列に比べてスパースである。

提案手法の F 値が他のどの手法よりも高かった楽曲 RWC-MDB-P-2001 No.5 におけるパラメタ β の条件ごとの目的関数 Proposed, SparseCount, Entropy の値のグラフを図 6 に示す。なお、描画する際には目的関数間の大

表 5 RWC-MDB-P-2001 No.7 ($\beta = 2.5$) のクラスタ間の遷移確率行列 (表 6 のクラスタとは対応していない)

クラスタ	a	b	c	d	e	f
a	0.36	0	0.09	0	0.09	0.45
b	0.15	0.77	0	0	0	0.08
c	0	0	0	0	0	1
d	0	0	0	0.11	0.11	0.78
e	0	0	0	0	0	1
f	0.05	0.03	0.07	0.08	0	0.78

表 6 RWC-MDB-P-2001 No.7 ($\beta = 1.5$) のクラスタ間の遷移確率行列 (表 5 のクラスタとは対応していない)

クラスタ	a	b	c	d	e	f
a	0.14	0.29	0.43	0.14	0	0
b	0	0.31	0	0.36	0.33	0
c	0	0.45	0	0.18	0.36	0
d	0.06	0.21	0.12	0.15	0.38	0.09
e	0.04	0.22	0.08	0.24	0.39	0.02
f	0.17	0	0	0.08	0.08	0.67

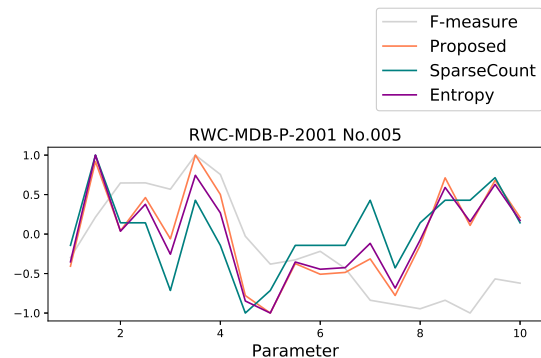


図 6 Proposed, SparseCount, Entropy の比較 (RWC-MDB-P-2001 No.5)

小関係に着目するため、それぞれ最大値が 1.0、最小値が -1.0 となるように正規化している。目的関数 Proposed は $\beta = 3.5$ の時に最大値をとり、F 値もそのときに最大値をとる。一方、SparseCount と Entropy は Proposed と類似した変動をするが、 $\beta = 1.5$ の時に最大値をとるため、Proposed に比べて F 値は下がっている。

5. 考察

目的関数 Proposed を用いて楽曲ごとに適切な HPSS のパラメタを求めることで、パラメタの最適化を行わない場合及びドラムパターンの遷移を考慮しない場合と比べて高い F 値が得ることができた。実際に目的関数 Proposed の値が最大値をとるところでは、クラスタリングの精度も高くなり、各ドラムパターンの遷移確率も遷移する先が限られたドラムパターンであるような遷移確率になっていた。目的関数 Proposed と類似の尺度である Entropy や SparseCount を用いた場合の F 値もパラメタの最適化を行わない場合およびドラムパターンの遷移を考慮しない場

合と比べて高くなった。これらより、ドラムパターンの遷移に基づいて設計した目的関数は、大域的なドラムパターン構造を信号処理に反映する目的に有用であったことがわかる。

一方でドラムパターンのクラスタリングの精度が低く、期待通りのF値の改善が見られない場合もあった。目的関数の最大化によって得られるクラスタリングの精度が低い場合は大きく分けて二つ考えられる。一つ目は、クラスタリングの精度と目的関数の値が連動しない場合である。これは記号処理のレイヤーの問題であり、目的関数に組み込まれた楽曲が有すると仮定した楽曲構造と、実際の楽曲の有する楽曲構造が異なっている場合に生じる。二つ目は、パラメタの変更がクラスタリングに影響を及ぼさない場合である。これは音響信号処理のレイヤーの問題であり、目的関数によって適切に最適化されていてもHPSSのパラメタの変化がクラスタリングの識別面に影響を及ぼさない場合に生じる。これに対しては最適化するパラメタを増やすなどの対応を行う必要がある。

また、ある程度正しくクラスタリングされないと適切な遷移確率が求められないため、不適切な遷移確率に基づいた情報、あるいは無意味な情報をフィードバックしてしまうという問題があることが分かった。k-meansクラスタリングには初期値依存問題があり、フィルインのパターンが適切に一つのクラスタとして識別されるかどうかは、初期値に大きく依存する。新しく別の方法によってフィルインを推定し、フィルインのパターンが初期値として選ばれやすくなる枠組みを構築する必要がある。

6. おわりに

記号処理によって得られる楽曲内の大域的なドラムパターン構造を、信号処理によるドラムパターンの認識へとフィードバックする枠組みを構築し、調波音・非調波音分離のパラメタを最適化することで、ドラムパターンのクラスタリング精度を向上させることができた。

今後の課題は記号処理を行う前のクラスタリングにおける評価関数を定義し、信号処理の結果を記号処理の推論にフィードバックする枠組みを構築する。例えば音響信号のレベルでのフィルインらしさを定義することができれば、k-meansクラスタリングの初期値に組み込むことでドラムパターンの認識の精度が上がるだけでなく、大域的な構造を考慮したドラムパターンの差し替えシステムにおける適切なドラムパターンの対応付けにも資する。これは信号処理と記号処理の相互作用フレームワークの一つとなる。

謝辞

本研究の一部はJST ACCEL (JPMJAC1602), JSPS 科研費 (16H01744, 16K12560) の助成を受けたものです。

参考文献

- [1] Bergstra, J. and Bengio, Y.: Random search for hyperparameter optimization, *Journal of Machine Learning Research*, Vol. 13, pp. 281–305 (2012).
- [2] Dixon, S.: Onset detection revisited, *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-6)*, Vol. 120, Citeseer, pp. 133–137 (2006).
- [3] Driedger, J., Müller, M. and Disch, S.: Extending Harmonic-Percussive Separation of Audio Signals, *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR)*, pp. 611–616 (2014).
- [4] Fitzgerald, D.: Harmonic/Percussive Separation using Median Filtering, *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx-10)* (2010).
- [5] Goto, M., Hashiguchi, H., Nishimura, T. and Oka, R.: RWC Music Database: Popular, Classical, and Jazz Music Databases, *Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR)*, pp. 287–288 (2002).
- [6] Jolliffe, I.: Principal component analysis, *International encyclopedia of statistical science*, Springer, pp. 1094–1096 (2011).
- [7] Oore, S., Simon, I., Dieleman, S. and Eck, D.: Learning to Create Piano Performances, *NIPS 2017 Workshop on Machine Learning for Creativity and Design* (2017).
- [8] Paulus, J. and Klapuri, A.: Music structure analysis by finding repeated parts, *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*, ACM, pp. 59–68 (2006).
- [9] Paulus, J., Müller, M. and Klapuri, A.: State of the Art Report: Audio-Based Music Structure Analysis., *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR)*, pp. 625–636 (2010).
- [10] Raphael, C.: Harmonic analysis with probabilistic graphical models, *Proceedings of the 4th International Society for Music Information Retrieval Conference (ISMIR)*, pp. 177–181 (2003).
- [11] Tanaka, T., Nishimoto, T., Ono, N. and Sagayama, S.: Automatic music composition based on counterpoint and imitation using stochastic models, *Proceedings of 7th Sound and Music Computing Conference* (2010).
- [12] Thompson, W., Sundberg, J., Friberg, A. and Frydén, L.: Rules for expression in the performance of melodies, *Psychology of Music*, Vol. 17, No. 1, pp. 63–82 (1989).
- [13] Widmer, G.: Learning expressive performance: The structure-level approach, *Journal of New Music Research*, Vol. 25, No. 2, pp. 179–205 (1996).
- [14] Widmer, G.: Machine discoveries: A few simple, robust local expression principles, *Journal of New Music Research*, Vol. 31, No. 1, pp. 37–50 (2002).
- [15] 浜中雅俊, 平田圭二, 東条 敏: 音楽理論 GTTM に基づくグルーピング構造獲得システム, *情報処理学会論文誌*, Vol. 48, No. 1, pp. 284–299 (2007).
- [16] 平田圭二, 青柳龍也: パーピープン: ジャズ和音を生成する創作支援ツール, *情報処理学会論文誌*, Vol. 42, No. 3, pp. 633–641 (2001).
- [17] 澤田 隼, 深山 覚, 後藤真孝, 平田圭二: 大域的構造を考慮したドラムパターンの差し替えシステム, *研究報告音楽情報科学 (MUS)*, Vol. 2018-MUS-119, No. 43, pp. 1–7 (2018).