

ハウスミュージックの楽曲構成の分析

ウルフ ジャスティン智也^{1,a)} 名越 崇晃^{1,b)} 北原 鉄朗^{1,c)}

概要: 本稿では、ハウスミュージック (以下ハウス) における楽曲構成の自動推定を目標とし、ハウスの楽曲構成を比較的単純な方法で分析した結果について述べる。既存の音源分離技術を用いてドラムス、ベース、ヴォーカル、その他に分離し、それぞれに対して音圧 (波形の二乗平均平方根) を計算してグラフとして描画した。また、分離前の音源の音圧に基づいてイントロ (導入部分)、ドロップ (盛り上がり部分)、ブレイク (盛り上がり部でない部分)、アウトロ (終結部分) に分けて色付けした。分析の結果、ドラムスとベースは楽曲構成の展開に与える影響が大きく、その他のパートは楽曲構成の展開には関与が薄いことが分かった。また、イントロとアウトロの判定に関して改善の余地があることが分かった。

1. 序論

ハウスミュージック (以下ハウス) はダンスミュージックの一種であり、ポピュラー音楽とは異なる構成で作曲される。いわゆる「A メロ」「B メロ」「サビ」といったメロディやコード進行を中心とした構成によってメリハリをつけるのではなく、各パートが同じパターンを何度も繰り返す中、各楽器パートが演奏を始めるタイミングを巧みに制御することで起承転結を形作る。

このような各パートの出現状況を、既存の楽曲から自動で抽出して可視化できれば、ハウスという音楽ジャンルを深く理解するきっかけを与えることができ、ハウスを作曲する際や、ハウス楽曲を扱った DJ プレイの際の参考資料にすることもできる。そこで、本研究では、様々なハウス楽曲の音源を自動で分析し、楽曲構成およびそれを形作る各パートの出現状況を可視化することで、ハウスを作曲したい人や、ハウス楽曲を扱う DJ に対して有用な情報を提供することを、最終目標とする。

楽曲構成の分析手法の提案は多々ある。例えば、自己類似度行列を用いた楽曲構成の可視化 [1] や、確率と探索アルゴリズムを用いた楽曲のパート分割 [2] といったものがある。しかし、ハウスに特化した手法はまだ提案されていない。

本稿では、ハウスの既存楽曲に対して各楽器パートの演奏内容を分析し、楽曲構成を自動推定することを目標とし、比較的単純な方法で楽曲を分析した結果について述べる。

既存の音源分離技術を使ってドラムス、ベース、ヴォーカル、その他の音源に分離し、各音源の音圧をグラフ化し、そこから読み取れる特徴を考察する。

2. 分析方法

ハウスを始めとするダンスミュージックにおける楽曲構成の呼称に明確な定義は存在しないが、本稿では以下の 4 つのセクションを採用する。

- イントロ: 曲の導入部分
- ドロップ: 曲における盛り上がり部分
- ブレイク: 曲の盛り上がり部分以外の部分
- アウトロ: 曲の終結部分

ドロップは、ポピュラー音楽における「サビ」に相当し、ブレイクは、サビの後の間奏やサビの前の A メロ・B メロに近い位置づけである。ハウスなどの音楽では、上でも述べたように、メロディやコード進行を変えるのではなく、新たなパートを加えたりして楽曲のメリハリを作ることが多い。そのため、ここで述べたような楽曲構成は、音圧に現れる可能性が高いといえる。

また、近年の音源分離技術の発展は目覚ましく、ミックス済みの音源からドラムス、ベース、ヴォーカルなどの各パート音源を高精度に取り出すことができる。そこで、各パートの音源を取り出し、パート別の音圧についても合わせて分析する。

2.1 音源の前処理

各パートの音源を音源分離システム Demucs[3] を用いて生成する。Demucs は、Meta の研究部門である Meta Research が開発を行っているオープンソースの音源分離

¹ 日本大学 文理学部 情報科学科

^{a)} wulf@kthrlab.jp

^{b)} nagoshi@kthrlab.jp

^{c)} kitahara@kthrlab.jp

モデルである。Demucs は、音源分離コンテスト「Music Demixing Challenge 2021」で優勝していることから、精度の高い音源分離が期待できる。Demucs は、入力として任意の音源ファイルを渡すと、4つのパート（ドラムス、ベース、ヴォーカル、その他）に音源を分離して出力する。本稿の実験では、入力する音源の形式は mp3 とし、出力される音源の形式は mp3、ビットレートは 320kbps、サンプリング周波数は 44.1kHz とする。

実行の際にオプションを付与する事で、6つのパート（ドラムス、ベース、ヴォーカル、ピアノ、ギター、その他）に分離することも可能だが、ピアノの分離の精度が必ずしも十分とは言えないことと、楽曲構成の分析においては4パートで十分と考えたことから、本稿の実験では4パートへの分離を行う。

2.2 各パートおよび分離前の音源に対する音圧の計算

音圧は、対象とする音源の波形の二乗平均平方根 (RMS) として計算する。具体的に、波形 $y(t)$ における音圧 $S_i(t)$ は下の式で表される。

$$S_i(t) = \sqrt{\frac{1}{T} \sum_{\tau=0}^T y(t+\tau)^2 dt}$$

上の式において、 i はパートを表す ($i \in \{\text{drums, bass, vocals, others, mixed}\}$)。RMS を計算する際の窓幅 T を 65000 サンプル、時間分解能 ($S(t)$ の時間間隔) を 16250 サンプルとした。

RMS は、Python の音楽分析用モジュール Librosa を利用して算出する。その後、未分離音源の音圧を基準として、最大値が 1、最小値が 0 となるように各パートの音圧の値を正規化する。

2.3 グラフの描画

各パート（ドラムス、ベース、ヴォーカル、その他）の音源から求めた音圧の時間変化 $S_{\text{bass}}, S_{\text{drums}}, S_{\text{vocals}}, S_{\text{others}}$ に加え、未分離音源の音圧 S_{mixed} に基づいて楽曲を4つのセクション（イントロ、ドロップ、ブレイク、アウトロ）に分類して色付けて描画する。各セクションの判定基準は以下の通りとする。

- イントロ (黄)：曲の開始時刻から、最初に S_{mixed} が 0.85 を超える時刻までの区間
- ドロップ (赤)： S_{mixed} が 0.85 を超えている区間
- ブレイク (緑)： S_{mixed} が 0.85 を下回っている区間（イントロ、アウトロに該当する区間を除く）
- アウトロ (青)： S_{mixed} が 0.85 を超える最後の時刻から、曲の終了時刻までの区間

この基準は、サビに相当するドロップは、基本的に他のセクションよりも音圧が大きいという仮説に基づく。0.85



図 1 Selecao - Mark Knight, Shovell

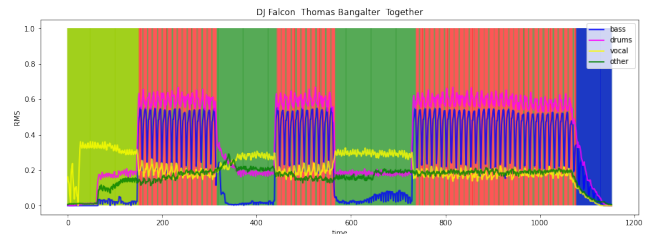


図 2 Together - Thomas Bangalter, DJ Falcon

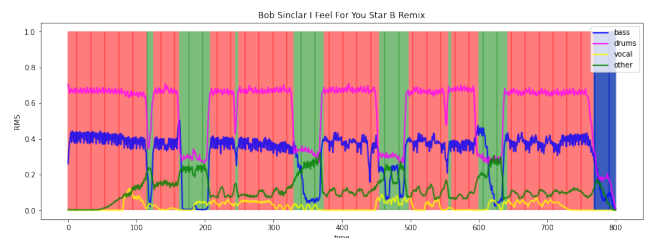


図 3 I Feel For You Star B Extended Remix - Bob Sinclar, Riva Starr, Mark Broom, Star B

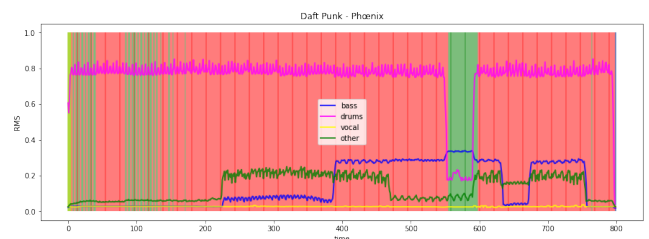


図 4 Phoenix - Daft Punk

というしきい値は、いくつかの楽曲を試しながら実験的に決定した。

3. 分析結果・考察

2. で述べた方法に基づいて、ハウスの楽曲を分析した。分析に用いた楽曲は、次の4曲である。

- (1) Selecao - Mark Knight, Shovell
- (2) Together - Thomas Bangalter, DJ Falcon
- (3) I Feel For You Star B Extended Remix - Bob Sinclar, Riva Starr, Mark Broom, Star B
- (4) Phoenix - Daft Punk

分析結果を図 1～図 4 に示す。

図 1 と図 2 の楽曲に関しては、4つのセクションが適切に分かれており、また、セクションが切り替わる場面では、

各パートの増減がある事も読み取れる。例えば図1の楽曲では、イントロから最初のドロップに切り替わる際に、ドラムスとベースの音圧が高まっている。また、楽曲の中盤では、ドロップが終わったと同時にドラムスとベースが離脱し、ヴォーカルが入り、次のドロップで再度ヴォーカルが離脱し、ドラムスとベースが入るという入れ替わりがあることが読み取れる。図1と図2は概ね意図通りに色分けが機能しており、典型的なハウスの例であるといえる。

一方、図3の楽曲はイントロ部分が、図4の楽曲はイントロ部分に加えアウトロ部分が消失している。特に図4の楽曲は、一部を除いて高音圧のドラムスが鳴り続けており、ドラムスの音圧だけが曲全体の音圧に大きな影響を及ぼしている。このことから、イントロ、アウトロ部分の自動推定をおこなうためには、各楽器の音圧に基づく判定のみでは不十分であることが分かる。

また今回使用した4つの楽曲すべてにおいて、ドロップ区間では必ずドラムスとベースが鳴っており、この2つのパートが音圧に大きく影響していることが分かる。また、ブレイク区間は大半でベースが鳴っておらず、ドロップからブレイクへの切り替えを演出する際に低音が抜ける傾向があることが分かる。

また、ドラムスとベースがセクションの切り替わりに関与していることが多いのに対し、ヴォーカルとその他はセクションの切り替わりに関与していることが少ない。このことから、ハウスの楽曲構成においてはドラムスとベースが展開を左右する役割を果たしていることが分かる。

4. おわりに

本稿では、ハウスの楽曲構成の理解を助ける分析の一環として、各楽器の音圧の描画および未分離音源の音圧に基づくセクション（イントロ、ドロップ、ブレイク、アウトロ）の分類・色付けを試行した。

実験から、ハウスの楽曲における各パートが構成の展開に与える影響が分かった。例えば、ドロップに入る際はドラムスとベースの音圧が高まるパターンが多い。反対に、ドラムスとベース以外のパートは楽曲構成の展開にそこまで関与しておらず、ハウスにおいては主にドラムスとベースが楽曲構成の展開における役割を果たしていることが分かった。

また、実験を通じて現状の自動推定の問題点が浮かび上がった。例えば、ドロップとブレイクの自動推定は音圧を参照する方法が効果的であったのに対して、イントロとアウトロは音圧が高い場合や低い場合が存在するため、音圧に基づく判定のみでは不十分であることが分かった。今後は、ドラムパターンやベース音の自動採譜技術、繰り返しパターンの検出技術など、様々な技術を併用して精度を改善していきたい。

謝辞 本研究は、科学研究費補助金 22H03711, 21H03572

の支援を受けた。

参考文献

- [1] Jonathan Foote: Visualizing Music and Audio using Self-Similarity, Proceedings of the Seventh ACM International Conference on Multimedia (Part 1), pp. 77–80 (1999).
- [2] Jouni Paulus, and Anssi Klapuri: Music Structure Analysis Using a Probabilistic Fitness Measure and a Greedy Search Algorithm, IEEE Transactions on Audio, Speech, and Language Processing, Vol. 17, Issue: 6, pp. 1159–1170 (2009).
- [3] Alexandre Défossez: Hybrid Spectrogram and Waveform Source Separation: ISMIR 2021 MDX Workshop, p. 11 (2021).