

# 音楽解析に基づく自動化 VJ システムの開発 — Ableton Live プラグインを用いた演奏連動型映像生成

コウ ウケツ<sup>†1</sup> 久原泰雄<sup>†1</sup>

**概要:** 本研究は、DJ と VJ がそれぞれ独立して成立してきた従来のライブパフォーマンス構造を背景に、DJ でも簡単に利用できる自動化型 VJ プラグインの開発を目的とする。一般的に VJ は事前に用意された映像素材を操作する形式と、TouchDesigner / Max などアルゴリズム生成を行う形式に分かれる。一方 DJ も完成楽曲を切り替える方式と、DAW のループ素材を組み合わせる即興的に構成する方式が存在する。しかし、DJ パフォーマンスは高い集中力を要求するため、音楽操作と同時に複雑な映像演出を行うことは容易ではない。また、VJ プログラミングの知識を持たない DJ や音楽制作者にとって、既存の VJ ソフトウェアの使用は難しい。これらの背景から、本プラグインは自動化と即興演奏への応答性を重視し、DJ や音楽制作者たちライブパフォーマンス中にも簡単に利用できる Ableton Live プラグインを開発した。

## 1. はじめに

一般的に、VJ パフォーマンスはおおよそ二種類に分けられる。一つは、Resolume などの専用ソフトウェアを用い、事前に準備した VJ 用映像を VJ パフォーマーがリアルタイムに操作、再生する形式である。もう一つは、TouchDesigner や Max などのソフトウェアを用い、パフォーマンス中に映像を変化させるか、あるいは DJ の演奏内容に応じて自動生成する形式である。後者には、常にアルゴリズムによる粒子表現や 3D モデルのダイナミックな変形などが含まれる。

そして、現在の DJ パフォーマンスにもおおよそ二種類に分けられる。一つは、Serato DJ や Rekordbox などの DJ ソフトウェアを用いて事前に用意した楽曲を切り替える形式である。使用される音源はほとんどが完成された楽曲である。この場合、音楽間の切り替え（ミックス）のロジックは基本的に「一曲から次の一曲へ」という構造を持つ。もう一つは、Ableton Live などの DAW ソフトウェアにおける Clip/Pattern モードを用い、事前に準備したループ段落素材を切り替えて演奏を構成する形式である。この形式では、必ずしも「一曲」という概念を必要とせず、複数のループを組み合わせてもパフォーマンスを構築する。また DAW 上であるため、即興的な楽器演奏をより簡単に加えることができ、結果として全体が一つの長い即興的“楽曲”として成立する場合もある。二番目の場合には、最近の電子音楽制作者たちはライビングをパフォーマンスする時、よく見られる。

一般的には DJ と VJ は別の役割であり、ライブパフォーマンスには両者の協力によって成立している。相対的に DJ パフォーマンスは高度な集中力が要求される。例えば T テクノのような強いビートを持つ音楽を扱う際に、集中力

を欠いてミックスのタイミングを誤ると重大なパフォーマンス上の失敗となる。観客がダンスに没入している環境では、そのズレがすぐ知覚されてしまうからである。そのため、DJ が自ら VJ 映像の切り替えまで同時に行う事例は多くない。

一方で、VJ が主導するパフォーマンス形態も存在する。この場合、VJ は高度な映像表現の知識を持ち、事前に自動化された映像切り替えアルゴリズムなどを構築することがよく見られる、そうすれば DJ と VJ を兼任しながらパフォーマンスを成立させる。しかし、VJ の学習やプログラミングに精通していない一般的な DJ にとって、このアプローチは簡単ではない。そのため DJ が VJ と協力する形態が広く見られるのである。

以上の背景から、自動化を中心に、DJ や音楽制作者でも簡単に利用できる VJ プラグインを開発するという考えが生まれた。プラグイン形式であるため DAW 上で動作することは自然であり、即興性が求められるライブパフォーマンスへの適応も重要である。そのため、本研究では即興演奏に応答可能な機能をプラグインに実装し、DJ が通常の音楽パフォーマンスの流れを保ちながら視覚効果を操作できるようにする。これにより DJ は音楽表現に没入したまま、視覚的パフォーマンスは本プラグインが自動的に処理する環境を実現することを目的とする。

## 2. 機能

本プラグインは、前述の背景を踏まえ、DJ や音楽制作者が高度な VJ パフォーマンスを容易に実現できることを目的として設計されている。そのため、操作性と即時性を重視しつつ、音響解析に基づく自動化処理および複数の映像表現モジュールを統合した構成となっている。本節では、本プラグインが備える主要な機能について詳述する。

<sup>†1</sup> 東京工芸大学大学院芸術学研究所

## 2.1 高品質な VJ 映像エフェクトのプリセット化

本プラグインには、粒子表現、幾何学変形、スペクトル可視化、音量に応じたモーション表現など、VJ パフォーマンスで広く用いられる複数の映像エフェクトがあらかじめプリセットとして搭載されている。これらのエフェクトは、Jitter を用いて構築されており、一般的なアルゴリズム VJ と同等の視覚的完成度を持つことを目標として制作したものである。ユーザはこれらのプリセットを選択する或いは自動的に切り替え、専門的な映像生成技術を必要とせずに、十分に魅力的な視覚表現を実現できる。

## 2.2 音響特徴量を基盤とした自動 VJ 効果切り替え機能

楽曲中の音響情報をリアルタイムで解析し、その特徴に基づいて映像表現を自動的に切り替える機能を備えている。具体的には、入力されたオーディオ信号から以下の指標を抽出する：振幅（音量）→ 映像の明度・サイズ・モーション量と連動。周波数帯域のエネルギー（Low/Mid/High）→ 色相変化やフィルタリング効果へ反映。オンセット（ビート検出）→ カット、フラッシュ、パターン切り替えなどのタイミング制御。

## 2.3 ユーザによる手動選択

自動化スイッチを操作すると、ユーザが手動で映像モードを選択・変更できるインターフェイスが提供されている。本プラグインでは、音楽制作者が慣れ親しんだ「パッド」型 UI を採用し、Ableton Live の Drum Rack に近い操作感を実現した。ユーザは、演奏スタイル、曲調、ライブ構成に応じて、任意のタイミングで映像効果を遷移できる。特に、クライマックスやブレイクダウンなどの場面で、視覚変化を手動で演出することが可能であり、全体のパフォーマンス性を高める機能となっている。

## 2.4 即興演奏へのリアルタイム応答

本プラグインは、即興的な操作に反応して映像を変化させる機能を備えており、MIDI 入力に強く対応している。たとえば、キーボード演奏や MIDI パッドの入力に応じて：映像オブジェクトの生成／消滅、色の変化、幾何学変形の強度などをリアルタイムで変化させる。これにより、DAW 上でのライブセットや電子音楽パフォーマンスにおいて、音楽と映像の一体化した即興表現が可能となる。自動化だけでなく「演奏表現の拡張」として機能する点が、本プラグインの重要な特徴である。

## 2.5 Ableton Live や Max 8 における利用

本プラグインは Max for Live を基盤としているため、Ableton Live 上での使用が可能である。また、Max や RNBO を用いることで処理系の柔軟性が高く、将来的に VST、AU などの一般的なプラグインフォーマットに拡張することも実現可能である。さらに、プラグイン読み込み機能を持つ一部の DJ ソフトウェアでも利用でき、DJ と VJ を同一環境内で完結させる可能性を開くものである。

## 3. 開発

本プラグインの開発目的は、VJ に特化した映像表現を実現することであり、一番重要な機能としては十分に高品質な VJ 視覚映像である。少なくとも一般的なアルゴリズム VJ と同等の視覚的完成度を達成することが求められた。また、本プラグインを汎用的な DAW に利用されると考えから、開発環境として Max を採用した。

Max に搭載されている Jitter 機能は、多様で高度な VJ 映像表現の開発に十分対応できる性能を有しており、さらに Max は Max for Live を用いることで、Ableton Live 上で利用可能な amxd 形式のプラグインとして比較的簡単にプラグイン化できるという利点を備えている。さらに、Max の RNBO モジュールを用いることで、C++ や JavaScript 形式のコードとして出力することも可能であり、将来的に VST などの一般的なプラグイン形式への展開にも対応し得る基盤となる。以上の理由から、筆者は本プラグインの開発環境として Max (Max for Live) を選択した。

UI デザインに関しては、本プラグインの利用者は DJ や音楽制作者と想定し、操作性および親和性を重視した。必要最小限の操作ボタンに加え、音楽制作者にとって馴染みの深い「パッド」型 UI を導入することで、ユーザに心理的な導入障壁を与えないインターフェイス設計を目指した。また、プラグイン側の UI は音声および映像の制御に限定し、VJ 映像自体は別ウィンドウのフローティングビューに表示する設計とした。これにより、パフォーマンス時にプロジェクターや複数ディスプレイへ容易に出力する、配置する可能となり、実演環境での利便性を高めている（図 1 参照）。



図 1 プラグインの GUI デザイン

本プラグインは、中央に配置した二つの主要スイッチを境界として、音声制御エリアと VJ 映像制御エリアの二つのセクションに分けられている。

### 3.1 音声制御エリア

本エリアの機能は比較的単純であり、主電源スイッチ、音声ファイル選択ボタン、再生と停止ボタン、再生状態を示すインジケータ、および音量メータを構成される。音楽特徴の解析および関連する計算処理は、内部アルゴリズムによって自動的に行われるため、複雑な UI を必要としないデザインとなっている（図 2 参照）。



図2 音声制御エリア

### 3.2 VJ 映像制御エリア

本エリアでは、VJ 映像表現に関わる主要なパラメータを制御する機能を提供する。具体的には、映像切替の自動化オン・オフスイッチ、自動切替の速度調整ノブ、全体的な VJ 表現の強度（映像の明度や変化の激しさに相当）の制御ノブ、そして各種 VJ プリセットを選択するためのパッド型 UI を備えている（図3 参照）。



図3 VJ 映像制御エリア

本プラグインにおける特徴的要素として、特にこの自動化切替スイッチが重要な役割を果たしている。

本研究における機能開発は、以下の三つの要素に大別される。(1)VJ 映像表現の開発、(2)映像の自動切替機能の開発、(3)音楽制作者によるライブ即興演奏へ応答する機能の開発。

#### (1)VJ 映像表現の開発

本プラグインにおける VJ 映像表現の開発は、「高い視覚的完成度」と「音楽入力への即応性」の両立を目標として進められた。一般的なアルゴリズム VJ では、粒子システム、ノイズベースの形状変形、スペクトラム可視化などが多用されるが、これらはリアルタイム処理の負荷が大きく、また高度なパラメータ調整が要求される。本研究では、Max の Jitter を中心にこれらの手法を最適化し、DJ や音楽制作者でも扱いやすい構成へ落とし込むことを重視した。

VJ 映像部分の開発においては、主にノイズ粒子の生成・演算、シェーダの変更、および三次元モデルの制御の三つの技術要素を組み合わせることで構築した。i)ノイズ粒子演算。ノイズ粒子表現の生成には、Max の Jitter が提供する OpenGL 機能を活用した。特に、jit.gen および jit.gl.pix を中心に用いることで、粒子の軌跡、動作、減衰周期などの動的パラメータを計算し、時間的変化を伴う VJ 映像の構築を可能にした。ii)シェーダ (Shadder) の変更。シェー

ダ処理については、GLSL を用いて Max のデフォルトシェーダファイルを直接編集し、筆者自分の規則を追加した。これにより、特定の VJ 効果が発生した際に、変更後のシェーダを自動的に呼び出す仕組みを実装し、視覚表現の柔軟性と一貫性を高めている。iii)立体モデル制御。立体モデルの制御には、jit.gen を用いてモデルの頂点情報を動的に操作し、他の素材やエフェクトと組み合わせることで、複合的な立体映像表現を実現した（図4 参照）。

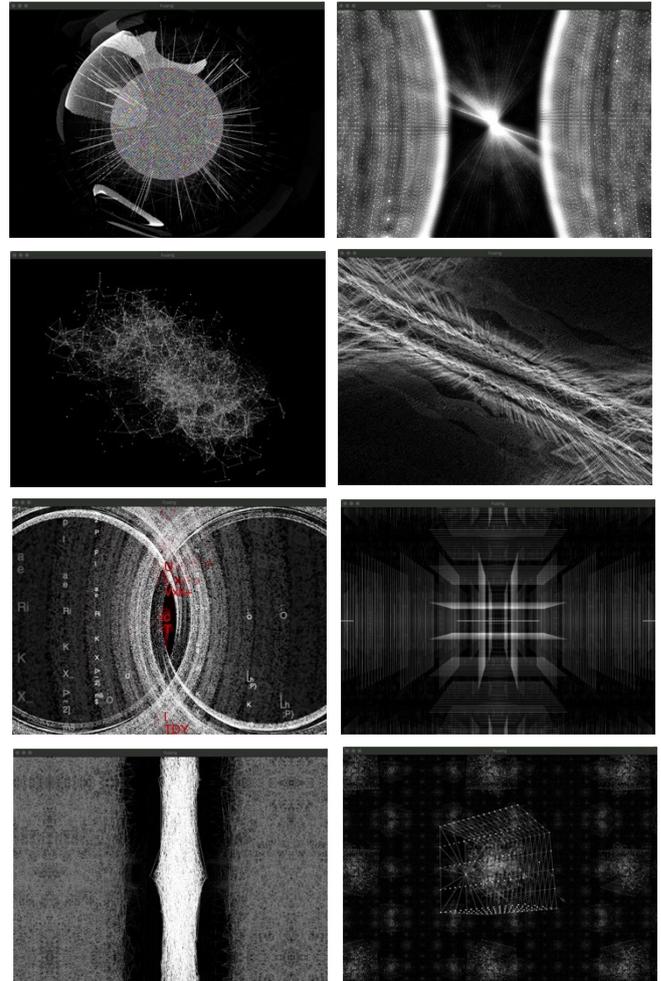


図4 立体映像表現

#### (2) 自動切替アルゴリズムの開発

本プラグインの主な利用者として DJ を想定した場合、多くの音源は強いリズムを特徴とし、特に低周波部分が密集する楽曲が多い。低周波部分の信号は干渉が大きく、成分が複雑に重なり合うため、正確に検出が困難である。そこで本研究では、周波範囲を限定するためのフィルタ処理を導入し、低周波成分を大幅に削減したうえで、open-hat や crash などの高域パーカッション (top 系サウンド) を主対象として検出する方法を採用した。

さらに本プラグインでは、ユーザ操作の low/high ノブを搭載し、フィルタのカットオフ周波数を任意に調整で

きるようにデザインした。これにより、曲のジャンルやミックスバランスに応じて、検出に最適な帯域をユーザが選択することができる。例えば、ハウスやテクノのようにハイハットが明確な楽曲では高域寄りに設定し、ドラムンベースのようにシャープなアタックが多い楽曲ではやや広めの帯域を確保するなど、柔軟な運用が可能である。

また、周波数帯域の変更は結果的に自動切り替えの「発生頻度」にも影響するため、ユーザは low/high ノブを操作するだけで、映像切り替えの密度をコントロールできる。この調整を楽曲テンポと組み合わせることで、テンポの倍数に基づく規則的な切替演出や、よりダイナミックで視覚的インパクトの強い演出効果を生み出すことができる。

以上のように、本研究の自動切替アルゴリズムは、DJ 現場での実用性を重視しつつ、楽曲の特徴を正確に反映したフィルタリング技術とユーザ操作性を統合することで、安定したリアルタイム映像切替を実現している。

### (3) ライブ即興演奏にตอบสนองする機能の開発

本機能では、DJ やライブパフォーマーが行うリアルタイムの演奏操作に対し、即時にตอบสนองする映像生成を実現することを目的としている。特に、外部デバイスから入力される MIDI 信号を基盤とし、その場で演奏される音高・和音構造を解析することで、VJ 映像が常に演奏の状態を反映するインタラクティブなシステムを構築した。Max の内蔵 JavaScript を利用して三和音（三つの構成音からなる和音）の種類を判定するアルゴリズムを実装し、和音の種類に応じて異なる視覚効果が発生するように設計した。

具体的には、三和音の判定結果に基づき、画面内に表示される立方体モデルの出現パターンや、映像全体の色彩変化が制御される。これにより、演奏者の即興的な音入力が直接視覚表現へと反映され、音楽と映像の一体感を高めることができる（図 5 参照）。

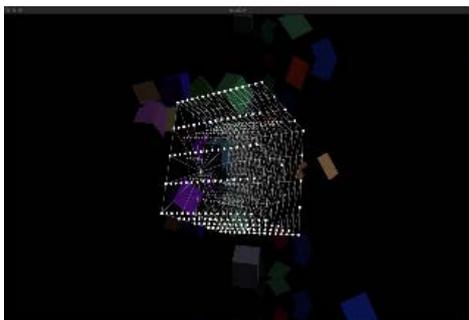


図 5 和音に応じた視覚効果

さらに、本システムは演奏者の操作に即時ตอบสนองするため、MIDI 入力の遅延を最小限に抑える工夫も行った。判別アルゴリズムは軽量化し、加えて映像生成処理も低負荷な構造に再設計することで、ライブ現場で求められるリアルタイム性を確保している。

具体の配色に関しては、人間の情動、色彩感、および和

音の相関関係に関する先行研究を主な参照としている。例えば、C メジャー (C Major) は多幸感を、C マイナー (C minor) は悲哀を感じさせるという性質がある[1]。音の高さ（ピッチ）と色彩の関連性が挙げられる。高音域は黄色や緑色、中音域は赤色や橙色、低音域は青色や紫色として知覚される傾向があり[2]、和音に対する聴覚と色彩に対する視覚の間には、一定の共感覚的効果（共感覚的対応）が存在すると指摘されている[3]。

しかしながら、本プラグインにおける具体的な色彩表現については、現時点では既存の文献や関連研究を開発初期の指針とするに留めている。本研究の最終的な目標は、本プラグインの公開テストを通じて、ユーザーデータを広範に収集し、そのフィードバックに基づいた実証的な修正を行うことにある。

その理由として、本プラグインは単なる色彩と和音の対応関係のみならず、図形の形状と心理的印象の相関性をも内包している点が挙げられる。例えば、円形と比較して三角形はより攻撃的な印象を与えるなど、形状によって惹起される感性は決定的に異なる。さらに、本プラグインはライブパフォーマンスでの使用を想定しており、身体の律動（リズム）と音楽の関係性についても、継続的な調査とデータ分析を要する[4]。

したがって、本稿では主に開発技術の側面に焦点を当てるものとし、その背景にある諸感覚の共鳴（シナジー）に関する詳細な考察については、今後の実証実験を通じたユーザーデータの収集・分析を待って、次報以降の課題としたい。

今後のバージョンでは、三和音だけでなく、四和音（セブンスコード）やテンションを含む複雑な和音構造の判定機能を追加する予定である。これにより、ジャズやフュージョンなど和音の多様性が高いジャンルでも柔軟に対応できる。また、音階（スケール）やモードの識別機能を実装し、演奏者が選択したスケールに応じて映像の色調や動きが変化する、より高度な音-映像インタラクションを実現する計画である。

## 4. 今後の展望

まず、3 和音以外の各種コード（5 和音、7 和音など）を識別するアルゴリズムを構築する。さらに、音階構造の判定機能も考えている。加えて、各種 VJ 表現のプロトタイプおよびコード識別アルゴリズムを C++ 環境へ移行し、JUICE フレームワークを用いた VST プラグインのアルゴリズム実装および UI デザインを行う予定である。

## 参考文献

- [1] 千葉 央輝, 榎松 理樹, 羽倉 淳, 藤田 ハミド. 音楽理論に基づく音声合成による感情表出手法. 全国大会講演論文集, 2010, vol. 72 (0), p.305-306.

- [2] Stephen E.Palmer ,Karen B.Schloss ,Zoe Xu , Lilia R.Prado-Leon .  
Music-color associations are mediated by emotion. PNAS, 2013,  
vol.110,no. 22, p.8836-8841.
- [3] 長田典子. 音を聴くと色が見える : 共感覚のクロスモダリティ.  
日本色彩学会誌,2010 ,vol.34 ,no.4 , p. 348-353.
- [4] 長嶋洋一.音楽の視覚化と身体性-メディアパフォーマンスに  
おける相互作用.芸術工学会誌,2011,Vol. 9, No. 1, pp. 23-30.