

FaceSession：顔を用いて創出される音楽での 観客参加型音楽の可能性

伏田 昌弘^{†1} 平林 真実^{†1} 小林 孝浩^{†1}

概要：近年、コンピュータをリアルタイムに操作して行う音楽ライブパフォーマンスが増えており、コンピュータの操作方法に焦点を当て、新しい演奏手法を提案する。顔の輪郭情報の抽出と感情（怒り）の AI 解析を組み合わせることで演奏手法の拡張を行う。感情の AI 解析を取り入れることで観客は自分の感情で演奏する新しい感覚を得ることができる。

1. はじめに

地方のアマチュアバンドシーンでは集客の問題を抱えている。本取り組みは、地方で低迷するライブイベントに活気を持たせることを狙い、「一回性の高い観客参加型の LIVE」を提唱するものである。地方のアマチュアバンドシーンでは、バンドの数が少なく、観客も同じである場合が多いため、イベントがワンパターンになりやすく、飽きられて足が遠のくということが多々ある。そこで本研究では、観客に繰り返し足を運んでもらうために、一回性を有した観客参加型の体験が有効であるとの仮説を立てた。一回性の担保、および、観客の演奏経験を補うために、現在利用可能なメディア技術を応用する。具体的には、誰もが有する表情を利用し顔の輪郭情報の抽出と感情（怒り）の AI 解析を組み合わせることで演奏手法の拡張を行う FaceSession というものである。本論文では、システムの実装と、これまでの実演を通じた改善を中心に報告する。

2. 作品について

2.1 設計

ライブの一回性に焦点を合わせ再現性の低い演奏となるシステムをつくる。顔認識というわかり易い仕組みを使い音楽経験がない人でも参加できる様にし、感情の AI 解析を用いて演奏を一回性のものにす。参加性を持たすことによって、観客は楽曲全体を聞きながら音のタイミングに合わせて顔を動かし、演奏に積極的に参加している感覚が得られる様にする。使用方法については、ライブの MC 時に観客を指名し演奏時は背景に模式的な顔の映像を投影する。表情の解析には 3 秒程度かかるので、曲全体の緩やかな効果に使用する。今回は、テンポへの反映および背景映像への反映へ使用する。

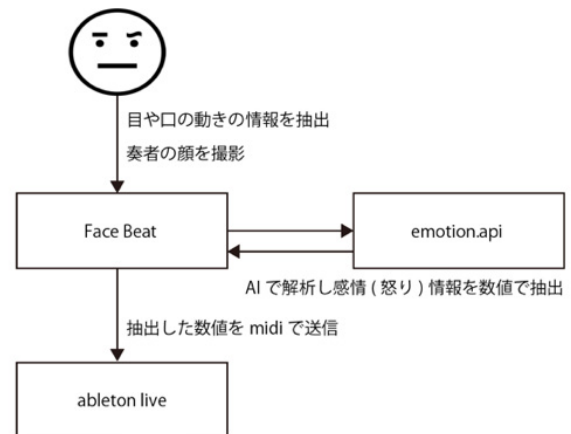


図1 システム概要図(Face Beat)

2.2 システム構成

openFrameworks(以下 oF)と node.js を用いて制作した。パソコンのカメラで得られた画像から顔を認識し目や口の動きの情報を抽出する。oF で奏者の顔を撮影し、表情を Microsoft の Emotion API(感情認識)を用い、AI で解析し感情情報を数値で抽出した。今回は、感情を数値化した値を音楽に使うことの有効性を評価する為にまず怒りのみを使用した。iphone 版もパソコン版と同様の機能で作成した。

FaceBeat ではドラムの演奏に、FaceBass では Bass の演奏に FaceSynth ではシンセサイザーの演奏に用いている。また、ライブで FaceSession 時、実際に PC に映っている顔の映像を演者の背後に出力する。それにより FaceSession に参加していない観客も Session の様子がわかる。

2.3 各楽器について

・ FaceBeat

- ① 口を開ける：CLAP の音が鳴る。
- ② 口を閉じる：ドラムのキックの音が鳴る。
- ③ 目を大きくする：ボリュームが大きくなる。
- ④ 目を小さくする：ボリュームが小さくなる。
- ⑤ 怒った顔をする：テンポが上がる。

表示されている顔が笑顔から般若になる。

[†] 情報科学芸術大学院大学 メディア表現研究科

・FaceBass

- ① 口を開ける：Bass の音が鳴る.
- ② 口を閉じる：Bass の音が消える.
- ③ 目を大きくする：音が高くなる.
- ④ 目を小さくする：音が低くなる.
- ⑤ 怒った顔をする：テンポが上がる.

表示されている顔が笑顔から般若になる.

・FaceSynth (iOS 版)

- ① 口を開ける：Synth の音が鳴る.
- ② 口を閉じる：Synth の音が消える.
- ③ 目を大きくする：ボリュームが大きくなる.
- ④ 目を小さくする：ボリュームが小さくなる.
- ⑤ 怒った顔をする：テンポが上がる.

表示されている顔が青から赤になる.

2.4 FaceSession

以上の三つの楽器 (Face Beat, Face Bass, Face Synth) を使用し,3人の奏者の目や口の動きと表情でセッションを行う.各奏者は,自分以外の音を聞きながら演奏することで,まるで本当のバンドセッションに参加しているかのような体験を得ることができる.

3. 先行事例・関連研究

21_21 DESIGN SIGHT で開催された,「AUDIO ARCHITECTURE: 音のアーキテクチャ展」で展示された自撮りと音楽ビデオが融合する AR アプリ「JIDO-RHYTHM」[1].この作品で鑑賞者は,自撮りの映像と,音楽を聴くという行為が重なる体験により,今までにない感覚が味わえる.この作品は自撮りした顔を映像作品の中に使っているのに対し,本研究ではカメラに写っている顔から音をリアルに生成している点に差分がある.

また,鑑賞者自身の顔を音符として表示し演奏するインタラクティブ作品「KAON」[2].この作品は,カメラに映る顔の位置を音に変えているが,本研究では顔の表情で演奏している点が違う.観客参加型の先行研究としては,音楽会場において演奏を行うパフォーマンスと観客のインタラクションを実現するためのシステム Cryptone [3].

この事例に対する本研究の特徴としては,ライブイベントでの音に対するアプローチをしている点である.

4. 実演

これまでに学内で組んだ f&u というユニットで 5 回のライブイベントでの実演,2回の展示を行ってきた.

(f&u は,ableton live という DAW を使ってリアルタイムに作ったトラックの上で

ベースとウィンドシンセを生演奏するユニットである.)

以下が詳細な日程と実施会場である.

08/04 IAMAS プチ・OPEN HOUSE 2018

08/27 ハボン de オープンマイク vol.22@名古屋 K・D ハボン

09/02 Interim Report edition3 @circus Tokyo

10/31 ハボン de オープンマイク vol.23@名古屋 K・D ハボン

11/02 NxPC.Live in 養老@養老公園

12/01 Ogaki Mini Maker Faire 2018

12/02 NxPC.Live vol.33 in OMMF 2018 @ソフトピアジャパンセンター

ライブで FaceSession 時,実際に PC に映っている顔の映像を演者の背後に,出力した.

そしてライブを終える度に課題を洗い出しアップデートを繰り返してきた.

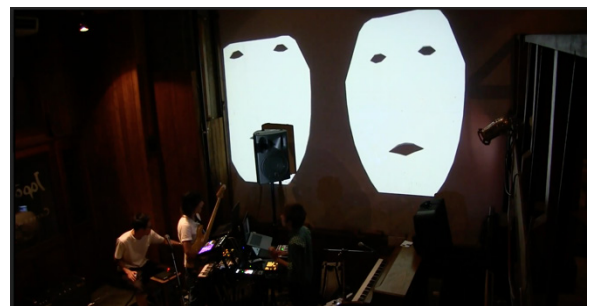


図2 ライブ風景 1(名古屋 K・D ハボン)



図3 ライブ風景 2(ソフトピアジャパンセンター)

4.1 課題の洗い出しとアップデート

「08/27 ハボン de オープンマイク vol. 22@名古屋 K・D ハボン」(図2)で最初の Live を行った.15 分間の Live パフォーマンスで,中盤の MC で FaceSession の説明を行い,演者 2 名でパフォーマンスを行った.観客からは新しい表現に対する興味や驚きの反応をもらった.特に AI を使って演者の顔の表情から感情を解析し音と顔に反映しているところに興味を持っていた.

だがこの Live で以下の課題がみつかった.

- (1) 顔のビジュアルが質素である.
- (2) モバイル版の Face 楽器がないので観客に体験してもらえない.

上記の課題をクリアし,「10/31 ハボン de オープンマイ

ク vol.23@名古屋 K・D ハボン」で観客参加型の Live を行った。

15 分間の Live パフォーマンスの中盤の MC で FaceSession の説明を行い、演者 2 名と指名した観客 1 名の 3 人で 2 分程度 FaceSession を行った。終了後、声をかけて感想を聞くと参加者からは、顔を動かすだけで音楽的な素養がなく演奏に参加できるので良かったという感想をもらった。他の観客からも、やってみたいという感想をもらった。

ここで、以下の機能アップが必要だとわかった。

- (3) 演者が使用している 3 台の Face 楽器のテンポが ableton live と同期していないので、FaceSession から次の曲に繋ぐ時にずれる。

上記の課題をクリアし「12/02 NxPC.Live vol.33 in OMMF 2018 @ソフトピアジャパンセンター」(図 3)でパフォーマンスを行った。前回同様 15 分間のパフォーマンスの中盤で観客 1 名に参加してもらって FaceSession を行った。顔の動きだけで演奏でき音楽的にも成立する(テンポ、キーの楽曲との一致)様にシステムを組んでいるので一定の成功をおさめた。終了後、声をかけて感想を聞くと参加してもらった観客からも楽しかったという感想をもらった。

だが、ここで 2 つの課題が見つかった。

- (4) 会場のネットワークが弱かったので感情解析の応答遅延が発生し期待通りに感情が演奏に反映されない。
- (5) ステージの明かりが暗かったのでカメラによる顔認識がうまくいかないことがある。

(4)については、感情解析の仕組みを自前で作れば解決できるが、取得できる値の精度に不安がある。

(5)については、macbook pro の画面の明るさを最大にし、カメラが認識しやすい様にライトを追加する等の物理的な対策をとる必要がある。

5 まとめと今後の展望

顔を用いて創出される音楽での観客参加型音楽により、一回性の音楽体験を観客にしてもらい、ライブに繰り返し足を運んでももらえるというところまでの確認はできなかった。

FaceSession を体験してもらった観客からは、楽しかったという声をもらえ、体験はしていない観客からは、やってみたいという声をもらった。

今後、観客の声をアンケート等でもらえる仕組みを作り、システムをアップデートしライブを繰り返すことで、ライブに繰り返し足を運びたくなるかを定量的に評価してきたい。

参考文献

- [1] jido-rhythm <http://kitasenjudesign.com/jido-rhythm/>
- [2] 太田高志 田中潤 KAON(顔音): 顔認識を利用したインタラクティブ・コンテンツ 芸術科学会論文誌 Vol. 10, No. 3, pp. 148-156
- [3] 平林真実 清水基: Cryptone: 音楽会場におけるパフォーマンスと観客の相互インタラクションのためのシステム インタラクティブ発表, 2013