|  |
| --- |
| 配信ライブ中のヘッドバンギングによる一体感向上のための 動作推定手法の提案 |
|  |
| 二宮洸太 |
|  |
| **概要**：近年，様々なジャンルのアーティストによって音楽ライブが開催されており，観客は演奏に合わせてサイリウムを振る，ヘッドバンギングをするなど，アーティストや他の観客との一体感や非日常感などを楽しんでいる．その一方でライブの模様を配信する配信ライブも多く行われている．しかし，会場でのライブと比べ，アーティストや他のファンとの関わりや一体感が希薄化する問題がある．そこで，配信ライブ中の視聴者の一体感を向上させることを目的に，ライブ中に行われるヘッドバンギングを媒介とし，その動作を検知し，タイミングを視聴者間で共有するシステムを提案する．本稿ではまずスマートフォンのセンサデータを手掛かりとして，ヘッドバンギング動作の推定に関する検討を行った．具体的には，ヘッドバンギング中のセンサデータに関するデータセット構築を行い，機械学習によるヘッドバンギングの推定を行った．その結果，93.5%の精度で推定を行うことができた． |
|  |
| **キーワード**：ライブ，配信ライブ，ロック，ライブモーション，ヘッドバンギング，スマートフォン，センサ |

# はじめに

## 音楽ライブとその文化

近年，Apple MusicやSpotifyなどのストリーミングサービスの普及，YouTubeやニコニコ動画などの動画配信サービス上のミュージックビデオやライブ映像等の音楽コンテンツの増加によって，音楽を日常的に楽しむ機会が増えている．また，音楽の楽しみ方も多様化しており，ミュージックビデオやライブ映像，テレビ放送などを鑑賞することだけにとどまらず，カラオケで歌うことや，ライブ会場やフェス会場で不特定多数の人と生演奏の音楽体験を共有することなどもあげられる．その中でもライブは，ライブ・エンタテインメント調査委員会が実施した調査[1]によると2019年の観客動員数は5497万人，市場規模は4237億円と過去10年で最大となっており，注目されていることがわかる．

ライブは会場にそのアーティストのファンが集まっているという特殊な環境によって，一体感や空気感などを感じることができ，通常の音楽視聴では得られない音楽体験がある．アイドルのライブにおいては，光る棒状のデバイス（サイリウム，ペンライト）や好きなメンバーへのメッセージが書かれたうちわをライブ中に振る行為や，楽曲中の決まった区間で掛け声を叫ぶ「コール」といった文化が見られる．バンド，特にラウドロックやヘヴィメタルといった激しい楽曲の多いアーティストのライブでは，「ライブモーション」と呼ばれるライブ特有の行動がある．ライブモーションには観客同士が体をぶつけあう「モッシュ」や，観客たちが円状に走り回る「サークルピット」，4つ打ちのビートにあわせて独特のステップを踏む「ツーステップ」，楽曲に合わせて頭を振る「ヘッドバンギング（以下ヘドバン）」などがある．このようなライブ独特の行動や気持ちの高まった観客たちの空気感によって，ライブ空間でしか体験できない一体感や非日常感をアーティストとファンは共有し，ライブを楽しんでいる．

## 音楽ライブの生配信とその問題点

近年の音楽ライブ市場では，そのパフォーマンスをインターネット上で生配信する配信ライブが盛んになっている．これは会場に行かなくても手軽にライブを視聴できることや，アーカイブとしてライブ動画を残すことで一定の期間自由に視聴できることから需要が高まっている．またCOVID-19の流行により，会場に多くの人を集められない状況下でもライブを行う方法としてさらに脚光を浴びている．

しかし，会場で開催されるライブに足を運ぶときとは異なり，配信ライブで視聴者が行えることは限られている．YouTube LiveやSHOWROOMといったライブ配信プラットフォームでは，コメントや投げ銭といった機能が提供されている．コメントはライブ中に高まった気持ちや応援メッセージを，アーティストや他の視聴者に向けて発信することができる機能である．また，投げ銭は特定のアイテムなどを購入することで，アーティストを支援することができる仕組みである．これらの機能によって視聴者は自身の熱意を，アーティストや他の視聴者に発信することが可能であるが，会場で開催されるライブのように他の観客とライブモーションを行うことで得られる一体感を感じることは難しい．

その一方で，配信ライブ中にそれぞれの視聴者がひとりでライブモーションを行っている場合がある．あるロックバンドの配信ライブ中には「パソコンに向かってヘドバンしている」というツイートが散見されており，実際にライブ会場で行うヘドバンを配信ライブ中にひとりで行っていることがわかる．またアーティストによる配信ライブに関するツイートでは，視聴者に対して配信ライブ中にヘドバンを煽るツイートも存在する．しかし，既存のライブ配信システムが提供する機能では，他者が行っているヘドバンなどの動作を知ることができず，他者の存在が希薄となってしまうため，一体感を得ることは難しい．

これらの問題を解決するために，配信ライブを支援する研究は多く行われている．しかし，サイリウムの同期や手を振る・拍手などのジェスチャを音楽に反映するといった研究は多くされているものの，ライブモーションに着目した研究はあまり行われていない．

## 研究目的

本研究では，配信ライブにおける，視聴者が行うライブモーションを通じた一体感の向上を支援することを目的とする．本稿では特に配信ライブでも視聴者が行いやすいライブモーションであるヘドバンを対象とする．具体的には，配信ライブ中のヘドバン動作を検知し，リアルタイムに視聴者間で共有することで，配信ライブ中の視聴者間の一体感を向上させることを目指す．そのためには，視聴者のヘドバン動作を推定し，その動作を視覚や聴覚を介して他者と一緒にヘドバンしていると感じられる形で共有・伝達する必要がある．ここで，実際のライブで一緒にヘドバンをしていると感じるのはヘドバン中の頭の上げ始め，下げ始めのタイミングが一致していることである．そこで，この上げ始め・下げ始めのタイミングを機械学習によって推定し，共有することで，ヘドバンを介した一体感の向上を実現する．

したがって，本稿では配信ライブにおける一体感を向上させる手法の前段階として，スマートフォンのセンサ情報を用いてヘドバン動作を取得し，機械学習によって推定することで，配信ライブ中のヘドバン動作の検知する手法を提案する．

# 関連研究

## ライブに関する研究

ライブに関する研究はこれまで多く行われている．ライブが観客にもたらす効果を調べた研究として，中村ら[2]は，ライブ演奏と録音演奏の差異を明らかにするため，それぞれの演奏を聴き比べる実験を行った．その結果，ライブ演奏のほうが録音演奏に比べ，好き・良い・また聞きたいという項目が高い評価であることを明らかにした．またライブ演奏には，穏やかさや生き生きといった感情を増幅する効果もあることを明らかにした．Weinbergら[3]は，音楽との関わりによる生活満足度の違いを，主観的ウェルビーイング（SWB）の観点で調査した．その結果，ダンスや音楽イベントの参加者は，参加してない人に比べSWBのスコアが有意に高く，特に他者と一緒に歌ったり踊ったりした参加者はよりSWBスコアが高いことが明らかになった．Brownら[4]は，ライブに行きたい理由に関する調査を行い，ライブ参加者がライブの雰囲気や演奏が毎回変化するユニークな体験を求めていること，アーティストを生で見たいこと，ファンとアーティスト並びにファン同士のコミュニケーションやインタラクションを求めてライブに参加していることを明らかにした．Swarbrickら[5]は，あるアーティストのファンとファンではない人にライブと音源をそれぞれ体験・聴取させ，動きを比較する実験を行った．その結果ライブを体験させた場合，音源を聴取させた場合に比べて観客は活発に動くこと，特にファンの方がファンではない人よりも活発に動くことを明らかにした．また，ファンとファンではない人は曲によって動きが異なる事例が見られ，ファン独自の盛り上がり方があることも明らかにした．ライブ中の観客の動きに関して，Silverbergら[6]は，ヘヴィメタルコンサート中の観客の動き，特に「モッシュ」という観客同士が体をぶつけあう動きに関する研究を行っている．ここでは，主体的にモッシュを行うActive Mosherと，主体的には行わず他者のモッシュによって巻き込まれるPassive Mosherに観客を分け，流体力学的にシミュレートすることで，観客の動きをモデル化している．これらのライブに関する研究から，ライブが音楽愛好者，特に特定のアーティストのファンにとって大事なものであることがわかる．本研究では，他の観客の存在の希薄化などにより，配信ライブで低減してしまうライブの一体感を，ヘドバン動作を共有することで向上させることを目的としており，配信ライブでの体験が向上することで，これらライブによる人への好影響が配信ライブでも得られることが期待される．

## ライブ支援に関する研究

ライブの支援を目的とした研究も数多く行われている．Freemanら[7]は，大人数の観客がそれぞれのスマートフォンを利用し，演奏者とインタラクションを可能とするmassMobileを開発した．これは観客らの投票によって楽曲のテンポやダイナミクスが変化するTeamWorkと，観客が描いた図形が映像として演奏者の背後に投影されるSketchingの2つの機能により，観客のライブへの参加感を高めている．Hirabayashiら[8][9]は，人間には聴こえない高可聴域音を利用し，各周波数にIDを割り当て，それを認識する技術（USC）を用いてライブ体験を変化させる研究を行っている．Cryptone[8]は，USCによってDJが観客のスマートフォンに提示される映像を制御するシステムであり，演者と観客のインタラクションを実現している．Sense of Space[9]は，USCを用いて観客のスマートフォンから流れる音声を制御することによって，観客参加型で音楽をつくるシステムである．Katoら[10]は，音楽再生をインターネット経由で同期し，機器を制御する大規模音楽連動制御プラットフォームSongle Syncを開発した．これは音楽理解技術Songle[11]のデータを用いて音楽に連動した映像や照明効果を，スマートフォンやIoTデバイスといった複数の端末で同期提示することで，一体感のある演出を実現したものである．また，アイドルのライブによく用いられるサイリウムに着目し，これを利用したライブ支援の研究も多く存在する．小幡ら[12]は，サイリウムの振りによって演者の衣装に装着したLEDの点灯するパターンが変化するシステムを開発した．川元ら[13]は，スクリーンに投影した映像に合わせてサイリウムを振り，映像の動きにマッチした場合サイリウムを点灯させることで，観客の動きを誘導するシステムを提案している．高橋ら[14]は，演者のファンサービスがどのファンに向かって行われているかを特定するために，サイリウムにカメラを取り付け，顔認識やポーズ認識を行えるシステムを提案した．岩本ら[15]は，アイドルライブでの初心者支援を目的として，ライブ参加経験のある経験者の動きに合わせてサイリウムを振動させることによって，振るタイミングを初心者に提示するサイリウムを作成した．武井ら[16]は，ユーザのサイリウムの振りを誘導することを目的に，1つのサイリウムの振り方に合わせて複数のサイリウムの発光パターンが変化する仕組みを作成している． 平川ら[17]は，無線通信でサイリウムの発光を制御するシステムLumiConneを開発し，実証実験を行い，150名での動作を確認している．堂林ら[18]は，ファンのサイリウムの振りに合わせて，アイドルの装身具が変化することでファンとアイドルのインタラクションを実現するSync☆Idolを開発した．またKayaliら[19]は，ライブで用いられる観客参加型のインタラクティブシステムについて演奏者と観客がどう感じているか調査を行った．その中でシステムはアーティストが制御したうえで，観客によってパフォーマンスが変わるようなものが好ましいこと，システムは観客の体験にプラスに作用すること，アーティストと観客のコミュニケーションを技術的に補っていること，フィードバックが多いほど観客の臨場感が高まることを明らかにした．またamazarashi[20]やTHE ORAL CIGARETTES[21]は観客のスマートフォンの操作によって，ライブ演出や楽曲が変化するライブ[22][23]を行っており，システムによるライブ支援は実用化されている．

会場で行われるライブだけでなく，配信ライブを支援する研究も多く行われている．柴村ら[24]は，観客の動きをカメラ映像で捉え，OpenPose[25]で推定した姿勢情報からジェスチャを認識し，演奏者の映像に重畳することで一体感を演出するシステムを作成した．また寺内ら[26]は，演者の動きによってペンライトが振動し，ペンライトの振りによって演者の衣装の発光や舞台演出が変わるインタラクティブペンライトを用いて一体感を向上するライブ支援システムAffinity Liveを開発した．米澤ら[27]は，個人の生配信におけるカメラワークや照明効果を視聴者が制御するシステムを作成し，ライブ演出の向上を図っている．浜中ら[28]は，視聴者の頭部に角度・距離センサを取りつけたコンサートスコープヘッドホンを提案している．これは視聴者の動きに応じてプロジェクタで投影する映像や音響効果を変化させるため，視聴者が特に見たい部分に焦点を当てたコンサート視聴を可能とするシステムである．吉田ら[29]は，ライブ映像に対して，自他の身体動作を重畳することによって動画視聴時の一体感を高める研究を行っている．またVRを用いた研究も行われており，粕谷ら[30]は，3次元のライブ映像・音声を撮影し，それを制御することで，ライブ映像をインタラクティブにVRで配信するプラットフォームを提案している．Kanekoら[31]は，VR上でライブ会場を再現し，ジャンプやサイリウムを振るといったライブモーションをアバターに反映することにより，VR空間上で非日常感や一体感のあるライブ再現をするシステムKSA2を開発した．

これらの研究は配信ライブの視聴において得られる情報量を増やすことにより，ライブへの没入性を高めることを目指している．本研究は配信ライブでの一体感の向上を支援することを目的としており，これらの手法と組み合わせることで，配信ライブにおけるユーザ体験をさらに向上させることができると考えられる．

## ライブ中のヘドバンに関する研究

ヘドバンに関する研究として，Hudson[32]は，ヘドバンの動きやそのパターン，価値，社会的な意味について考察しており，そのルーツはアフリカン・アメリカン文化にあり，ヘドバンの始まりはヘヴィメタ黎明期にあると述べている．またWalser[33]は，ヘヴィメタルの歴史についてまとめた自著の本の中で，ヘドバンは音楽のビートにあわせて激しく頭を振ることであり，モッシュ・クラウドサーフといったライブモーションの中でもヘドバンは象徴的な役割を果たしていると述べている．USAのヘヴィメタルバンドであるMetallica[34]の研究を行ったPillsbury[35]は，スラッシュメタルの音楽は，ヘドバンやモッシュを通じて演奏者や視聴者をその雰囲気に飲み込ませることができると述べている．またPurcell[36]は，ヘドバンをバンドのプレイや演奏に没入する方法であると述べている．

ヘドバンを用いた研究として，Bardosら[37]は，ギター型コントローラで音源を選択し，ヘドバンによって音源の再生を認識することで，ミュージシャンでなくとも作曲できるシステムBangaramaを提案している．Merrill[38]は，ギター演奏時のエフェクトの変更を頭の動きのトラッキングと，ジェスチャ認識で行うシステムを開発した．Mollら[39]は，ヘドバンを撮影し，その映像をPoseNet[40][41]を用いて取得した骨格情報からヘドバンのリズムを推定し，ビートとヘドバンのマッチによって得点の得られる音楽ゲームHeadbangZを作成した．

このようにヘドバンはその歴史や人々への影響が研究されており，ゲームや演奏への応用もされている．本研究ではヘドバンを，これまで行われていなかった配信ライブに用いることで，視聴者の一体感を向上させることを目指す．

## センサ情報を利用したジェスチャ・モーション推定に関する研究

スマートフォンの普及に伴い，スマートフォンのセンサ情報を利用したジェスチャ・モーション・行動推定の研究は多く行われている．Peiら[42]は，センサ情報とGPSの位置情報を組み合わせ，人がいる環境（コンテキスト）を推定している．Anjumら[43]は，センサ情報を用いて人間の活動状態を7つに分類する機械学習モデルを構築し，それを用いて毎日のアクティビティログを作成できるAndroidアプリケーションを実装している．Leeら[44]は，長時間のスマートフォン利用による健康への悪影響を防ぐために，センサ情報から姿勢情報を推定し，各姿勢の合計時間に応じてユーザに通知するシステムを作成している．Nipunら[45]は，建設業に従事する人たちの作業関連筋骨格系障害（WMSDs）を防ぐために腕と腰に2台のスマートフォンを取り付け，そのセンサ情報から姿勢の推定を行った．この推定結果を用いることで現場の作業を評価し，WMSDsを防ぐための対策を講じることができる．

音楽に関する研究にもセンサ情報は使われており，安永ら[46]は，音楽からの舞踊動作の生成に音楽特徴量に加えて，ユーザの動きの盛り上がりを加速度センサとして特徴量化し，キャラクタの動きをインタラクティブに制御できるシステムを開発した．Kankeら[47]は，ドラムの中で使用頻度の低い打楽器の仮想化を目的として，ドラムスティックに加速度センサを取り付け，その情報を用いて仮想化された打楽器の叩打を推定している．

センサ情報を用いて人間の姿勢や行動を推定していることを踏まえ，本研究ではヘドバンの動きがセンサデータとして取得できると考え，機械学習を用いて推定を行う．

表1 データ収集に用いた楽曲

|  |  |  |
| --- | --- | --- |
| 曲名 | 歌唱アーティスト | 時間 |
| Twilight | Fear, and Loathing in Las Vegas[47] | 3:54 |
| Monolith | Crossfaith[48] | 3:43 |
| The Revelation | coldrain[49] | 4:22 |
| JACK. B | SiM[50] | 4:04 |
| T R A N S l a t e d | Survive Said The Prophet[51] | 3:42 |

# ヘドバンセンサデータセット構築実験

本研究の目的である，配信ライブにおけるヘドバンを介した視聴者の一体感向上を実現するために，視聴者の動きをセンシングし，ヘドバン動作をリアルタイムに推定する必要がある．そこでまず，データ収集システムを開発し，ヘドバンセンサデータセットの構築を行った．

## データセット構築設計

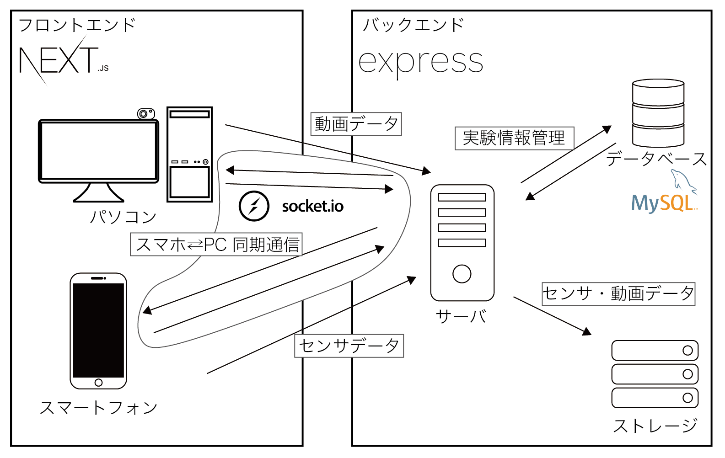
ヘドバンの推定は，配信ライブ中にリアルタイムで行うことを想定しているため，多くの視聴者が利用できる環境やデバイスで行うことが好ましい．そこで本研究では，パソコンでライブ映像を鑑賞しながら，ポケットに入れたスマートフォンを用いて，ヘドバン情報の収集を行う環境を想定する．このとき，動画から骨格情報を取得できるOpenPoseなどの手法を用いて，パソコンのカメラデータから，ヘドバン推定を行う方法も考えられるが，計算量が膨大であり，配信ライブのような環境で大人数の情報をリアルタイムに処理することは難しい．そのため，スマートフォンのセンサデータを用いて，ヘドバン情報の収集を行う．また，ヘドバンは音楽を聴きながら行うため，データセット構築においても実際のライブでヘドバンが行われる楽曲を聴きながらデータの収集を行う．

図1 データ収集システム概要図

今回使用する楽曲を表1に示す．これらは実際に著者がライブに行き，ヘドバンを行った曲の中で，ヘドバン時間が長いものを選定した．一方で，ヘドバンは視聴者が自身の判断で行うものであるため，データセット構築実験でヘドバンを行う区間に関しては実験協力者に任意で決めてもらうこととした．このとき，どこでヘドバンを行っているかをアノテーションする必要があるため，データ収集中の様子を動画で撮影する．過去の研究[48]でも，正解データの付与に動画を使用することが行われている．しかし，すべての動画に対して手動でヘドバン区間や周期をアノテーションするのは困難であるため，OpenPoseの姿勢情報を用いた自動化を併用して行う．これに関しては3.4節で詳しく述べる．

## データ収集システム

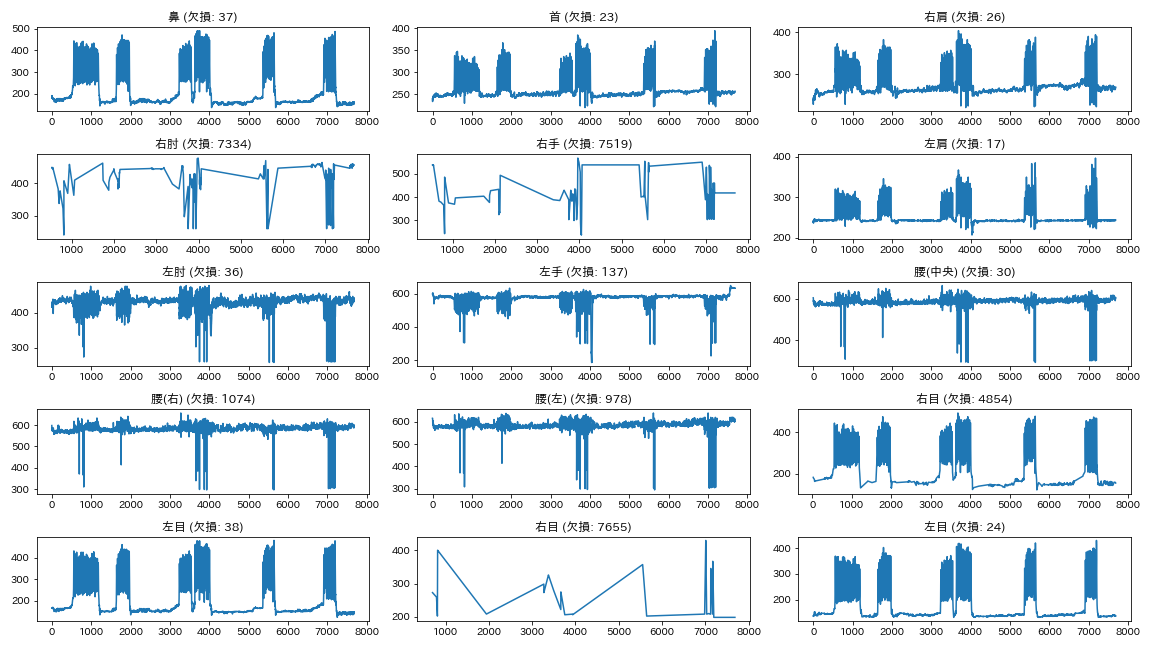
センサデータと動画データの対応を行うために，実験時に聴いてもらう楽曲に同期した状態でセンサデータと動画データの収集を行うWebシステムを開発した．システムの概要図を図1に示す．実装において，バックエンドはExpress.js，MySQL，フロントエンドはNext.jsを用いた．なお，パソコンとスマートフォン間の同期通信はsocket.ioを用いた．実験では，パソコンとスマートフォンそれぞれでシステムにアクセスを行う．スマートフォンで楽曲を選択し，データセット構築を開始すると，パソコンに開始時間が送られることで，スマートフォンとパソコンで時刻を合わせて同時にデータ収集が行われる仕組みとなっている．スマートフォンとパソコンは実験協力者が各自所有しているものを利用した．ただし実験システムの都合上スマートフォンはiOSに，パソコンのブラウザはGoogle Chromeに限定した．なおセンサ情報として3軸加速度（x， y， z）と3軸デバイス方向（alpha， beta， gamma）を60fps，動画は30fpsで収集した．

## データセット構築実験の実施

実験協力者は，ヘドバンを行うようなライブによく行く大学生6名（男性5名，女性1名，著者含む）である．実験協力者には事前にデータセット構築で使用する楽曲を聴きこんでもらい，どこでヘドバンを行うかについて各自の判断で決めてもらった．また，実験前に動作確認と実験の流れを把握してもらった．なお，実験は各実験協力者の自宅で行ってもらうこととした．これはセンサデータのみをデータセットとして利用することから，場所の影響を受けにくいと判断したためである．実験ではイヤホンで音楽を聴きながら，表1に示した楽曲に対し，それぞれ5回ずつヘドバンを行ってもらった．そのためデータセットは実験協力者6名×使用楽曲5曲×5回の150データが集まった．

ヘドバン動作の確認をわかりやすくするため，実験中はカメラに対し，横向きでヘドバンを行うように指示をした．またスマートフォン（iPhone）はイヤホンジャックを上，画面を肌と逆側，ポケットの位置を右ポケットにするよう指示し，向きや位置を統制した（図2参照）．さらにヘドバン時以外はジャンプ等の過度な動きは控えてもらった上で自由に音楽を聴くよう指示した．またヘドバンは首や腰に大きな負担がかかるため，首や腰に配慮した上で，複数日に分けて実施することを推奨した．

## OpenPoseによるヘドバン区間の自動付与

本稿では実験で得られたセンサデータに対して，実験で収集した動画データを手掛かりとして，ヘドバン区間とその周期に関する情報を付与する．ここで，ヘドバンには頭を横に振るものや全身を使わず首だけで行うものもあるが，本稿で扱うヘドバンは全身を使った上下に振るものに限定する．ラベルとしては非ヘドバン時を0，ヘドバン中の頭を上から下に振り下げる動作（振下動作）を1，ヘドバン中の頭を下から上に振り上げる動作（振上動作）を2と設定する．

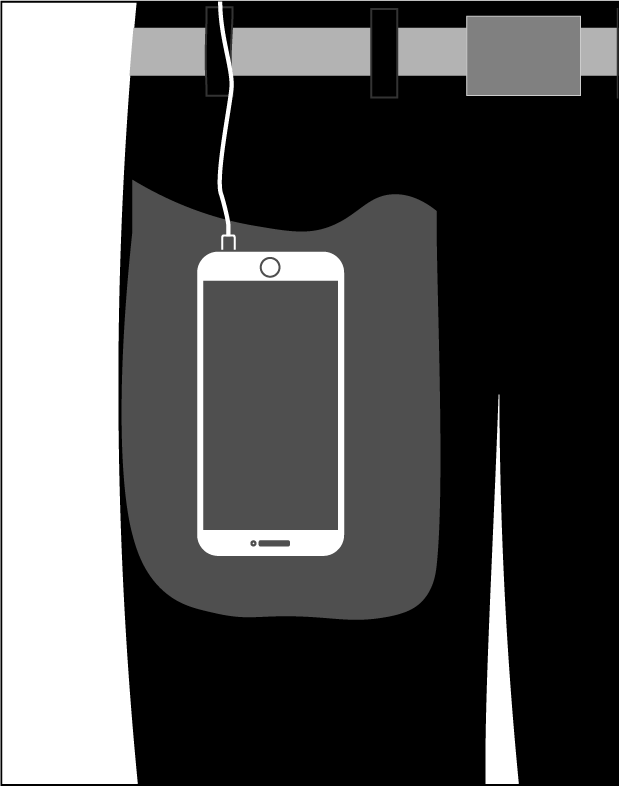
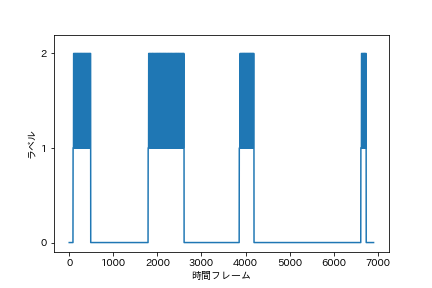
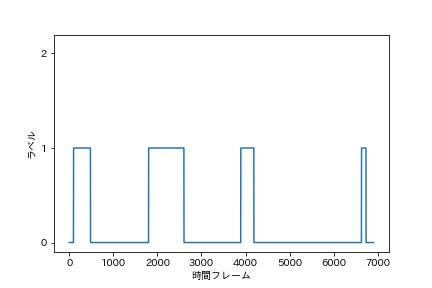
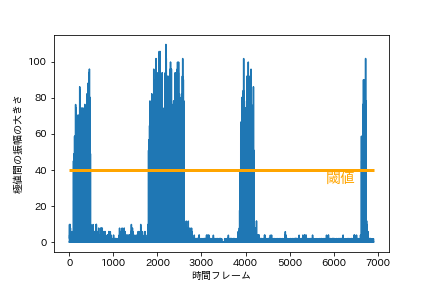
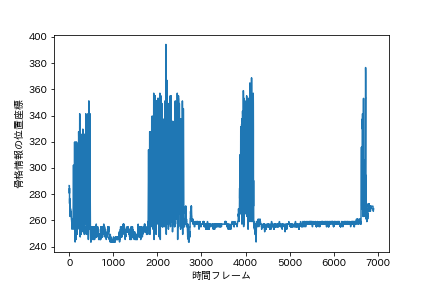
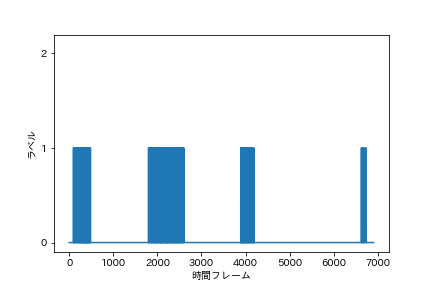
実験で得られた動画を人の姿勢推定モデルであるOpenPoseを用いて解析を行った．OpenPoseは動画像から人物の骨格情報を25点で抽出し，それぞれ2次元座標として出力することができる．実験で収集した動画に対して骨格情報の抽出を行ったところ，上半身，特に耳・目・肩といった部位でヘドバン中は大きく振動するという特徴が表れていることがわかった．そのうち上半身15点のy座標を時間軸上にプロットしたものを図3に示す．なお各部位の特徴の表れ方や欠損値の数は試行ごと異なるため，自動付与においては試行ごとに使用する部位を設定した．部位の選定は欠損値がなるべく少ないこと，図3と同様のグラフを全データで作成し，それを参照しヘドバン時と非ヘドバン時の値の違いがわかりやすいことを基準とした．ラベリングの過程を図4に示す．これはあるユーザの右肩のy座標の値を利用している．アルゴリズムで自動抽出可能なデータの例を図4(a)に示す．

図2 データ収集時のスマートフォンの状態

図3 あるユーザのOpenPoseによって抽出された骨格情報（x座標: 時間，y座標: y軸の位置座標）

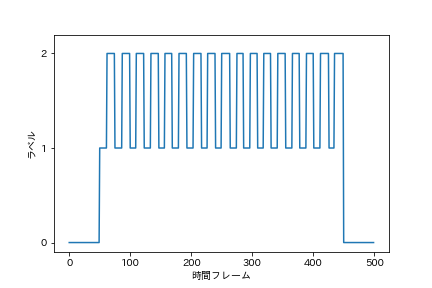


1. 付与対象の骨格情報
2. 振幅の大きさと閾値



1. 範囲ラベリング後
2. 手動ラベリング後
3. 極値ラベリング後

図4 ラベリング過程

ヘドバン動作の判定には動画を用いるが，動画を細部まで見てアノテーションするのは大変である．そこでヘドバンを行っている区間を自動抽出することで，アノテーションの易化を図る．ヘドバンは頭の振上動作と振下動作を繰り返す周期運動である．そこで，この周期性を利用して，周期の山と谷となる部分を抽出し，その間の振幅の大きさを求め一定の閾値でチェックすることで，ヘドバン区間を導出できるのではないかと考えた．具体的には，ヘドバン区間自動抽出アルゴリズムを構築し，ヘドバン区間の自動抽出を行った．

アルゴリズムとしては，はじめに山と谷となる部分を導出する．これは数学的には極値を求めることと同義であり，微分した値の符号が変化する点を求めることで，極値が算出される．次に極値間の振幅の大きさを計算し（図4(b)参照），これが閾値を超えればヘドバンとラベリングする（ラベルを1に設定する）．しかし，これではヘドバン時に頭の最も下がったときと最も上がったときのみ（極値のみ）しかラベリングされない（図4(c)参照）．そこで1のラベルが集中している区間を割り出し，その区間すべてを1とラベリングするようにした．その結果が図4(d)である．

このアルゴリズムを用いて，ヘドバン区間の自動抽出を行い，その区間に対し，振下動作，振上動作のラベル付与を手動で行った．なおアノテーションシステムはPythonを用いて実装した．その結果が図4(e)である．ヘドバンの判断については頭の位置に着目し，最も頭の高い位置から低い位置への移動区間を振下動作，最も低い位置から高い位置への移動区間を振上動作とした．またヘドバン中は振下動作と振上動作が交互に繰り返されるため，ラベルとしては図5のように1（振下動作）と2（振上動作）が交互になっている．

## 値の変換

機械学習による推定を行う上で，加速度，デバイス方向それぞれについて値の変換を行う．

図5 ラベリング区間の詳細

まず，加速度については，x，ｙ，ｚそれぞれ向きの値を合成した合成加速度を算出する．式としてはと表される．これにより，端末の傾き等による方向の変化の影響を除外する．

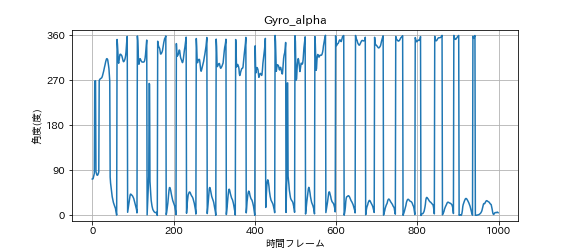
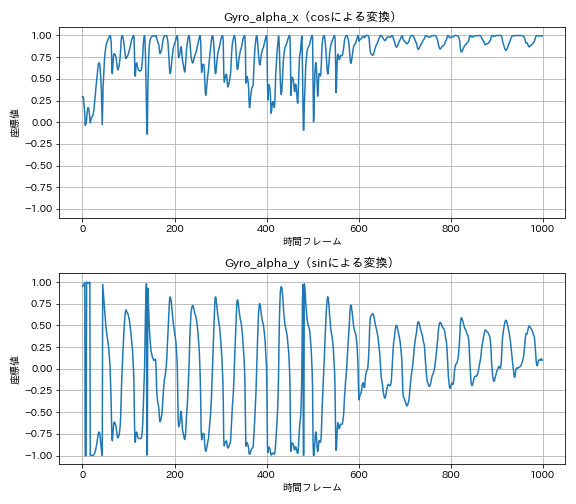
デバイス方向はalpha，beta，gammaの3種類を取得している．これは，それぞれz軸，x軸，y軸を中心としたデバイスの動きであり，角度として取得できる．このうち取得したalpha値の一部を図6に示す．alpha値の範囲は0～360°となっているため，361°以上となった場合は360°引かれた値が記録される．そのため360°から361°へ変化した場合，記録上は360°，1°と記録されてしまう．これではデバイスの動きを正しくとらえているとは言えない．そこで取得されたデバイス方向の角度をとしての形に変換し，2次元座標系にマッピングすることによって補正を行う．マッピング後の値を図7に示す．なおalpha以外のbeta，gammaについても同様に変換する．

図6 取得されたデバイス方向の値の一部

図7 変換後のデバイス方向の値のグラフ

# ヘドバンデータの分析

データセット構築時に得られたデータをもとに，ヘドバン中の振下開始，振上開始の推定に向けた分析を行う．ここでは，ヘドバン動作がセンサにどう表れているか，ヘドバン周期，ヘドバン区間の違いの3点について分析を行う．

## センサデータとヘドバン動作の対応に関する分析

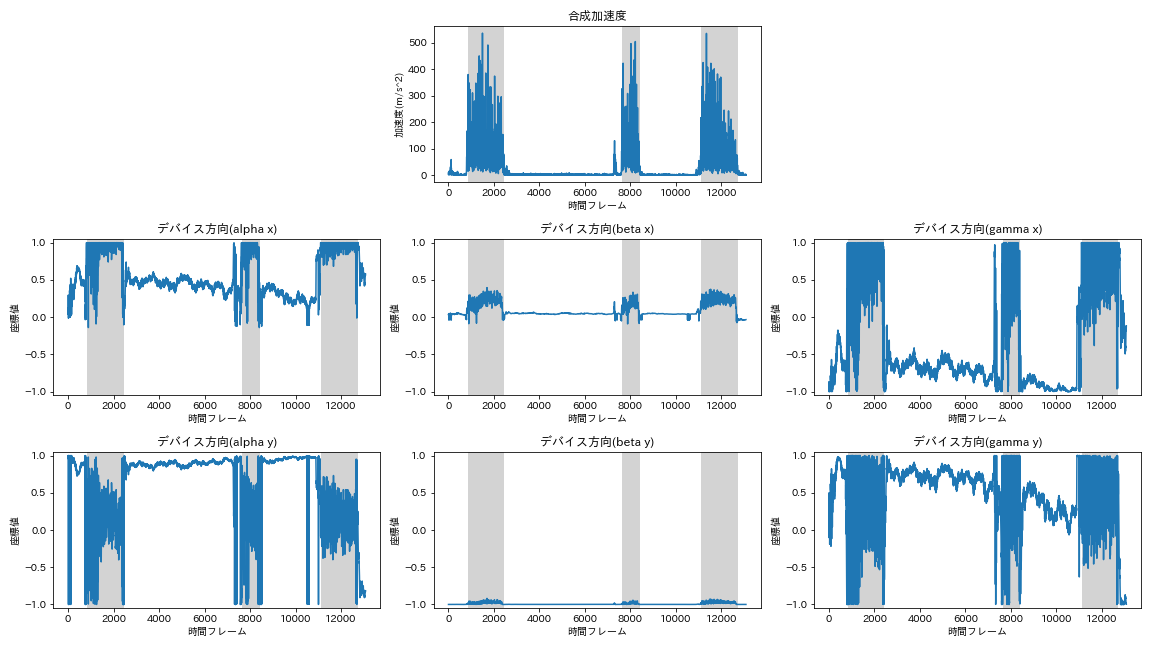


図8 センサの値とラベル区間の対応のグラフ

ヘドバン動作とセンサの関係について明らかにするために，センサデータとラベルの関係を可視化する．なおセンサは60fps，ラベルは30fpsであるため，時間によって位置を合わせた上で，センサ2フレームに対してラベル1フレームとしてプロットしている．あるユーザの1試行全体での様子を図8に示す．青色の線がセンサの値，背景色が灰色の部分はラベリングされたヘドバン区間を表している．図8より，合成加速度，デバイス方向ともに振動している部分とヘドバン箇所が対応していることがわかる．

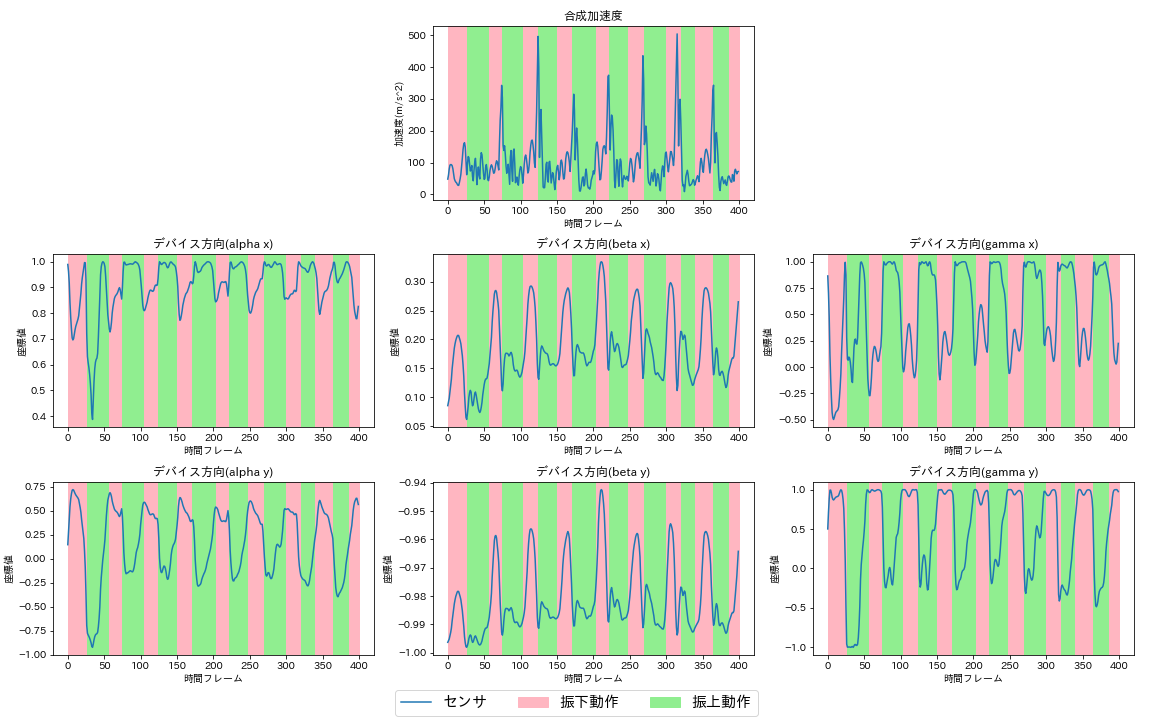


図9 センサの値と振下・振上動作ラベルの対応のグラフ

あるヘドバン区間に関して，振下動作，振上動作とセンサの対応をグラフ化したものを図9に示す．青色の線がセンサの値，赤色の区間が振下動作，緑色の区間が振上動作を表している．合成加速度，デバイス方向ともに周期的な波形をしており，ヘドバンの頭の動きの周期性がセンサに表れていると考えられる．また，振下動作と振上動作でそれぞれの区間でのグラフ形状に着目すると，それぞれのセンサ値において類似性が見られる．ここから，振下動作と振上動作にそれぞれ特徴があることがわかり，この特徴量からそれぞれの動作の推定が可能であると考えられる．

## ヘドバンの周期に関する分析

次にヘドバンの周期について分析を行う．今回推定したいのは振下・振上開始の推定である．特徴量抽出において，どの程度の時間でデータを切り取るかの検討には，実際のヘドバンがどのくらいの時間で行われているかを考慮する必要がある．そこで，ラベリングされた振下・振上動作がどのくらいで行われているかを分析する．

ユーザごとのヘドバン周期フレーム数の分布を図10に示す．グラフの青色が振上動作，オレンジ色が振下動作となっている．ユーザごとに見てみると，user1，user3，user4，user6は振上動作でのピークが12～13あたり，振下動作でのピークは9～12あたりに分布している．user2は振上動作では21～28あたりで広く分布しており，振下動作では19あたりでピークとなっている．user5はピークが2つあり，振上動作は25～26で最大，13で2つ目のピーク，振下動作では21を最大，13で2つ目のピークとなっている．これらのデータから，2倍周期でヘドバンを行っているユーザがいることと，どのユーザにおいても振上動作に比べ，振下動作の周期フレーム数が短いことがわかる．また，user2とuser5の最大のピークについて着目すると，user1，user3，user4，user6とuser5の2つ目のピークのフレーム数の約2倍になっている．これはヘドバンを2倍周期で行っているため起こったものと考えられる．

ここから振下動作，振上動作の特徴を正しく捉えるためには最低でも10フレームより小さいフレーム数で特徴量化を行う必要があると考えられる．

## ヘドバン区間に関する分析

今回のデータ収集実験では楽曲中のどこでヘドバンを行うかに関しては，ユーザ（実験協力者）の判断にゆだねている．そのため，ユーザ間でヘドバン区間が異なることが考えられる．そこで，ラベリングデータからヘドバンを行っている箇所を可視化し，分析を行う．ここで，ユーザは1曲につき5試行行っているため，この5試行の中でヘドバン区間が異なることがある．この場合は同じパターンが最も多い試行を可視化データとして採用する．

可視化した様子を図11に示す．なお動画収集条件の影響で，ユーザ間で数フレームずれが発生している場合がある．楽曲ごとユーザによる区間のばらつきを見ると「T R A N S l a t e d」はすべてのユーザが同じ区間でヘドバンを行っているが，他の楽曲ではばらつきがある．しかし，区間ごとに見ると多くの場合，複数のユーザがヘドバンを行っている．今回，実験対象はヘドバンを行うようなライブによく行く人に限定している．また実験前には楽曲を聴きこんでもらい，どこでヘドバンを行うかを決めてもらっている．これらを考慮すると，ヘドバンを行うようなライブに行く人たちは，楽曲の特性からどういう部分でヘドバンを行うかを暗黙的に理解していることが考えられる．しかし，楽曲のどのような特徴によって，ヘドバン区間を決めているかは明らかにされていないため，楽曲の音響特徴量等から分析する必要がある．

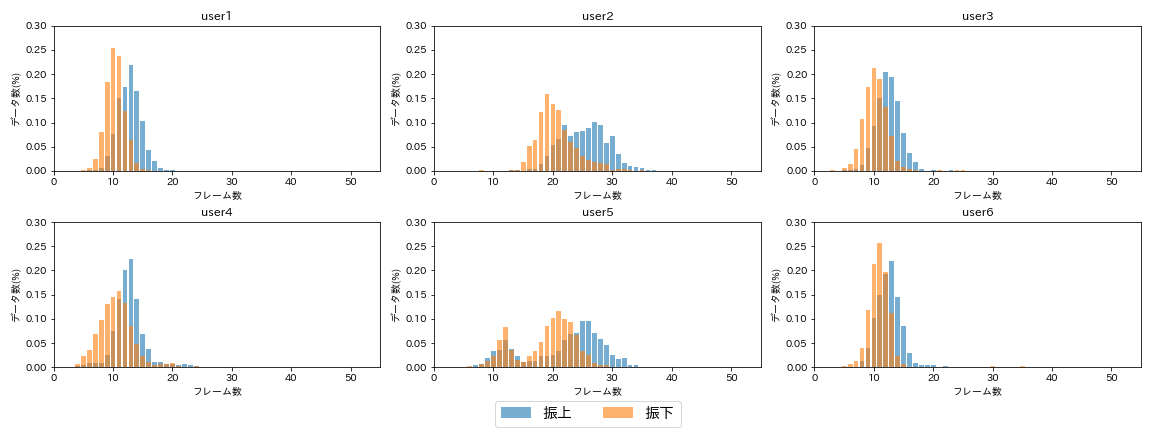


図10 ユーザごとのヘドバン周期フレーム数の分布

今回のデータ収集は実験協力者がそれぞれの家で，ひとりで行ったものである．しかし，実際のライブでは多くの人が集まっており，周りにあわせてヘドバンを行うため，ヘドバンを行う区間のばらつきは少ないものと考えられる．また，配信ライブにおいては，今回の実験環境と同様にそれぞれの家で，ひとりで行う環境となるためばらつきが生じる．ヘドバンの共有においてはこのばらつきを利用することで，ヘドバンしているか，していないかだけでなく，どのくらいヘドバンをしているかを共有することができる．これは実際のライブでは行えないことであり，配信ライブならではの体験を作り出せる可能性がある．

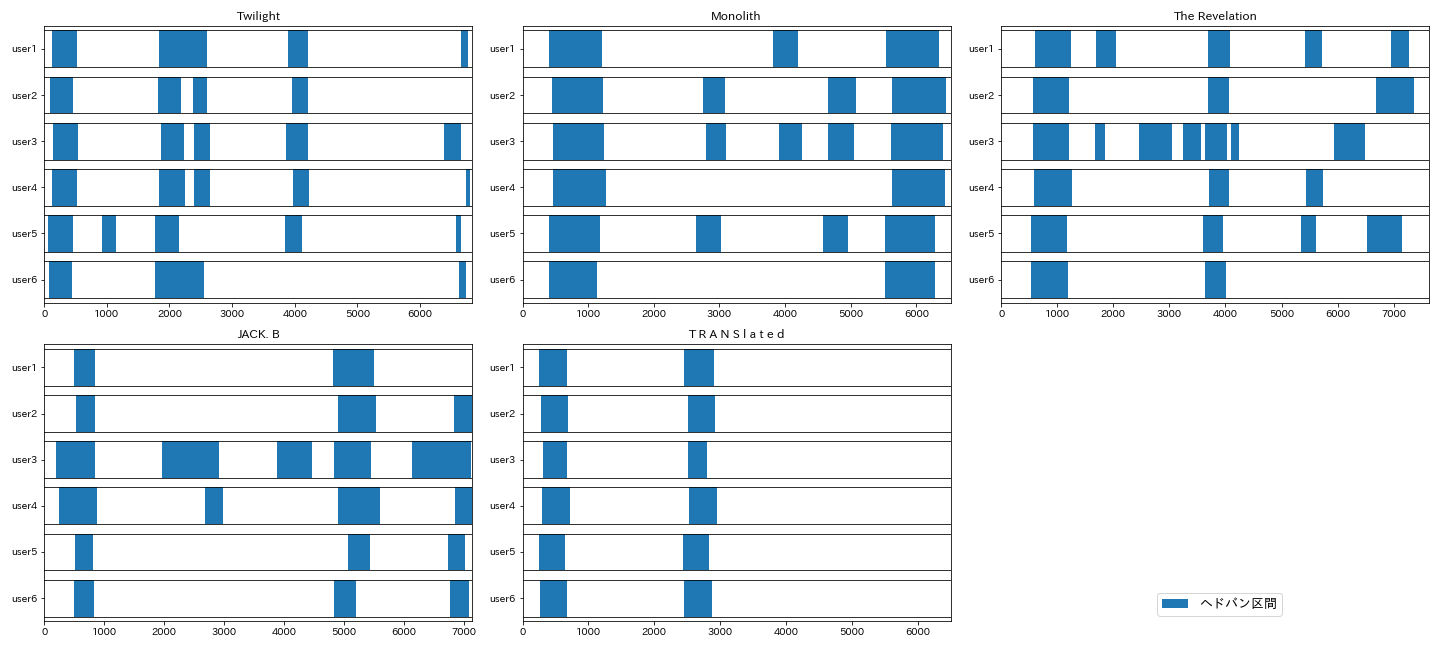


図11 楽曲ごとにユーザのヘドバン区間を可視化した様子

# 機械学習による推定

3章で構築したデータセットを用いてヘドバンの推定を行う．5.1節ではデータセットから学習する上での特徴量抽出の方法について述べる．5.2節では実際に学習を行い，その結果について考察する．5.3節では予測までの時間と特徴量の切り出す時間を変化させた際の，推定精度の変化について考察する．

## 特徴量抽出

ここでは，データセットからの特徴量抽出について述べる．今回構築したデータセットはラベルが30fps，センサデータが60fpsとなっている．そのため，フレームレートの違いに関しては，事前に時間基準でそろえた上で，ラベル1フレームに対してセンサ2フレームを切り取ることで整合させる．

まず，データの切り取り方について述べる．本研究では，配信ライブにおいて，ヘドバンを共有することを目的としている．ここで，実際のライブで一緒にヘドバンをしていると感じるのは，頭の下げ始め・上げ始めのタイミングが一致しているときである．そのため，タスク設定としては，頭の下げ始め，頭の上げ始め，その他の3値分類として行う．下げ始め・上げ始めに関しては各振下・振上区間の開始3フレーム目を開始地点とした．これはアノテーションと実際のヘドバンの様子の動画を見比べて決定した．ヘドバンを行っていない区間では，1秒ごとにデータをずらし，切り取り利用する．

また，配信ライブ中にリアルタイム推定を行うことを想定しているため，ある時点までのデータを用いてその先の予測を行う．その概略を図12に示す．ある推定したいフレームに対して予測するまでの時間となるインターバルを設け，その前のフレームを特徴量区間とし，このデータを用いて特徴量化を行う．インターバルと特徴量区間は，4.2節で振上・振下動作が10フレーム，時間にして0.33秒だったことから，それ以下にすることが望ましい．

次に特徴量化は，先述の特徴量区間内でデータを切り取り統計量化することで行う．データとしては合成加速度と3軸デバイス方向に加え，それぞれの差分データを利用する．差分データは隣接するデータをとすると，差分とあらわされる．これらのデータに対し，センサの種類ごとに，平均，標準偏差，最大，最小を求め，特徴量化する．そのためデータの次元としては，センサ7種×生データ・差分の2種×統計量4種の56次元となる．

## 学習と推定結果

5.1節の方法で特徴量化した特徴量を用いてモデルの構築を行う．機械学習のアルゴリズムはRandom Forestを用い，実装にはPythonの機械学習ライブラリであるscikit-learn[54]を用いた．本節ではインターバルを0.1秒，特徴量区間を0.3秒としてモデルを構築し，特徴量の検証を行う．このモデルをベースモデルと呼称する．なおインターバルと特徴量区間の時間と推定精度の関係に関しては，5.3節で検証する．3章で構築した150試行分のデータに対し，特徴量化を行い，86602データを得た．そのうち，80%にあたる69281データをTrainデータ，20%にあたる17321データをTestデータとしてランダムに分割した．

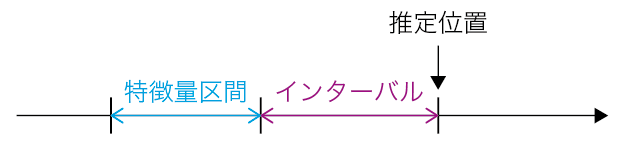


図12 推定位置と特徴量の関係

ベースモデルの精度を表2に示す．なお多クラス分類であるため，適合率，再現率，F値はマクロ平均をとったものである．正解率は93.5%，F値も0.912と高い精度で推定できていることがわかる．また，このモデルの混同行列を表3に示す．縦が正しいクラス，横が分類されたクラスである．今回のデータセットは非ヘドバンのデータが多い偏ったクラスである．それでも，混同行列，再現率の結果から振下開始，振上開始をうまく推定できていることがわかる．以上のことから，今回のモデルで高い精度で非ヘドバン，振下開始，振上開始を識別できることがわかった．

## 予測までの時間と推定精度

本稿ではリアルタイム推定を想定しているため，なるべく短い時間で推定することが求められる．しかし，時間が短いほど，特徴を捉えることが難しくなり，推定精度は低下すると考えられる．そこで，実際にインターバルと特徴量区間の時間を変化させ，推定精度への影響を検証する．特徴量は5.1節の方法で抽出し，インターバルと特徴量区間の時間を変化させる．インターバルは0.100，0.075，0.050秒の3種，特徴量区間は0.30，0.25，0.20，0.15，0.10秒の5種を採用する．機械学習アルゴリズムはRandom Forestを用いる．

各インターバルでの精度を，0.100秒は表4，0.075秒は表5，0.050秒は表6に示す．どのモデルにおいても特徴量区間の時間が短くなるほど精度が下がっている．特に再現率が下がっていることから，振下・振上開始の推定が難しくなっていることがわかる．これは特徴量として使える時間が短くなるために，特徴をうまく捉えられないことが原因と考えられる．次に同じ特徴量区間の時間内でインターバルの時間の違いによる変化を見る．どのインターバル時間においても，特徴量時間0.30秒のモデルでは正答率93%，0.10秒のモデルでは87%であり，インターバル時間の違いによる差異は見られない．これは，インターバル時間の変化が0.025秒と短いこと，インターバルのずれに応じて捉えられる特徴も変化していることなどが理由として考えられる．以上の結果から，実際の運用において時間を決定する場合は，インターバルの時間より，特徴量区間として利用する時間が重要であると考えられる．

配信ライブにおいてはユーザによって遅延が発生する．しかし共有においてはタイミングが一致することが重要であり，遅延が障害となる．今回インターバル時間による精度低下が見られなかったことを考量すると，遅延にあわせてインターバル時間を変化させて推定することによって，遅延による影響を除外できる可能性がある．

# リミテーション

表2 ベースモデルの精度

|  |  |  |  |
| --- | --- | --- | --- |
| 正解率 | 適合率 | 再現率 | F値 |
| 0.935 | 0.924 | 0.900 | 0.912 |

今回構築したモデルで，93.5%の精度でヘドバンの推定ができることが明らかになった．しかし，配信ライブでヘドバンの共有を行うためには課題が残されている．

まず，今回使用したモデルでは，振下・振上開始以外のヘドバン中のデータを含んでいないことが挙げられる．そのため，ヘドバン中の開始地点かそうでないのかの分類ができない．そこで，開始地点以外のデータを含んでモデル構築を行ったところ，開始地点の推定精度は約60%であった．しかし，この推定精度では正しく推定できているとは言えない．この問題の解決には，ヘドバン周期が役に立つと考える．ヘドバンの周期は著者の経験上，楽曲の区間に依存する．そのため，各楽曲のヘドバン区間とその周期を考慮して，ユーザの振下・振上開始を推定することで，高精度な推定が行えると考える．

また，配信ライブでは遅延が生じることが想定される．本稿の目指すシステムは，大人数の視聴者が同じタイミングでヘドバンを行い，それを体感することで一体感を発生させるものである．そのため，遅延によってタイミングがずれてしまうと一体感の低減や違和感につながってしまうと考えられる．5.3節の結果から，遅延具合にあわせてインターバルを変化させることでこの問題の解決ができると考えられる．しかし，インターバルの変化では大きな遅延に対応することは難しい．この場合は，遅延具合が近い視聴者間でヘドバン情報の共有を行うことが考えられる．どの程度遅延しているかは配信開始からの時間を記録する等の方法で取得可能であると考えられる．しかし，この手法では共有できる視聴者の人数の低減につながるため，どの程度の視聴者数がいれば一体感を感じられるかについても調査を行う必要がある．

表3 ベースモデルの混同行列

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | 予測結果 | | |
| 非ヘドバン | 振下開始 | 振上開始 |
| 実際の分類 | 非ヘドバン | 10661 | 120 | 100 |
| 振下開始 | 306 | 2759 | 152 |
| 振上開始 | 284 | 158 | 2781 |

表4 インターバル0.100秒でのモデルの精度

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 特徴量 区間 | 正解率 | 適合率 | 再現率 | F値 |
| 0.10 | 0.870 | 0.846 | 0.795 | 0.817 |
| 0.15 | 0.901 | 0.882 | 0.845 | 0.862 |
| 0.20 | 0.913 | 0.901 | 0.865 | 0.882 |
| 0.25 | 0.930 | 0.918 | 0.892 | 0.904 |
| 0.30 | 0.935 | 0.924 | 0.900 | 0.912 |

# まとめと今後の展望

本稿では，配信ライブ中の視聴者のヘドバン動作のリアルタイム事前推定に向けて，スマートフォンを用いて楽曲視聴中にヘドバン動作を行ってもらい，3軸加速度，3軸デバイス方向のセンサデータを収集する実験を行い，ヘドバンセンサデータセットを構築した．またRandom Forestを用いて機械学習による推定を行い，93.5%の精度でヘドバンの振下開始，振上開始，ヘドバンを行っていない状態の分類を行うことができた．

今後の展望として，6章で上げたヘドバン中の開始地点かそうでないかの分類を実現するとともに，人々のヘドバンの特性のさらなる解明，ヘドバン情報共有方法に関する検討を行う予定である．ヘドバンの特性の解明としては，4.3節で明らかになったユーザによるヘドバン区間の違い並びに，ヘドバンを行う区間の特徴を明らかにすることが挙げられる．ライブで観客の多くは暗黙的にヘドバンを行う区間を理解しているが，その楽曲的な特徴は明らかになっていないため，今後明らかにしていく．また楽曲からのヘドバン区間の推定が可能になれば，先述のヘドバン開始地点の分類が可能になる．また，さらに多くの人数のヘドバンデータを収集し，利用・分析を行うことで，モデル精度の向上や特性の解明を行っていく予定である．ヘドバン情報の共有・伝達に関しては，ヘドバン中に画面を見ることができないことを考慮する必要がある．配信ライブでは各個人がイヤホンを装着可能であること，センサデータ取得にスマートフォンをポケットに入れていることを考慮すると，ライブ音声を利用して聴覚情報として共有する方法や，スマートフォンを振動させ触覚情報として伝達する方法などが考えられる．どういった共有の仕方が一体感の向上を支援することが可能であるかについても，今後検討を行っていく．

表5 インターバル0.075秒でのモデルの精度

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 特徴量 区間 | 正解率 | 適合率 | 再現率 | F値 |
| 0.10 | 0.871 | 0.844 | 0.798 | 0.818 |
| 0.15 | 0.901 | 0.886 | 0.843 | 0.862 |
| 0.20 | 0.913 | 0.898 | 0.867 | 0.882 |
| 0.25 | 0.924 | 0.914 | 0.881 | 0.896 |
| 0.30 | 0.933 | 0.923 | 0.897 | 0.909 |

表6 インターバル0.050秒でのモデルの精度

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 特徴量 区間 | 正解率 | 適合率 | 再現率 | F値 |
| 0.10 | 0.870 | 0.844 | 0.795 | 0.816 |
| 0.15 | 0.900 | 0.883 | 0.841 | 0.860 |
| 0.20 | 0.912 | 0.896 | 0.865 | 0.879 |
| 0.25 | 0.921 | 0.911 | 0.878 | 0.893 |
| 0.30 | 0.933 | 0.920 | 0.897 | 0.908 |

**謝辞**

　本研究に際して，様々なご指導をいただきました中村聡史先生に深く感謝いたします．また，論文執筆や議論にご協力いただきました中村聡史研究室に先輩・後輩・同期，また快く実験に協力してくれた友人たちに感謝申し上げます．

**参考文献**

1. “2020 ライブ・エンタテインメント白書”. https://live-entertainment-whitepaper.jp/, (参照 2021-01-23).
2. 中村晃, 相良陽一郎. ライブ演奏と録音された演奏が聴衆者に与える印象の差異, 感情心理学研究, 2014, vol. 21, no. 2, pp. 72-79.
3. Weinberg, M. K., Joseph, D.. If you’re happy and you know it: Music engagement and subjective wellbeing, Psychology of Music, 2017, vol. 45, no. 2, pp. 257-267.
4. Brown, S. C., Knox, D.. Why go to pop concerts? The motivations behind live music attendance, Musicae Scientiae, 2017, vol. 21, no. 3, pp. 233-249.
5. Swarbrick, D., Bosnyak, D., Livingstone, S. R., Bansal, J., Marsh-Rollo, S., Woolhouse, M. H., Trainor, L. J.. How Live Music Moves Us: Head Movement Differences in Audiences to Live Versus Recorded Music, Frontiers in Psychology, 2019, vol. 9, pp. 2682.
6. Silverberg, J. L., Bierbaum, M., Sethna, J. P., Cohen, I.. Collective Motion of Moshers at Heavy Metal Concerts, Physical Review Letters, 2013, vol. 10, no. 22, pp. 228701.
7. Freeman, J., Xie, S., Tsuchiya, T., Shen, W., Chen, Y., Weitzner, N.. Using massMobile, a flexible, scalable, rapid prototyping audience participation framework, in large-scale live musical performances, Digital Creativity, 2015, vol. 26, no. 3-4, pp. 228-244.
8. Hirabayashi, M., Shimizu, M.. Cryptone: interaction between performers and audiences with inaudible DTMF sounds, Proc. of SIGGRAPH Asia 2012, 2012, pp. 1-4.
9. Hirabayashi, M., Eshima, K.. Sense of Space: The Audience Participation Music Performance with High-Frequency Sound ID, Proc. of NIME 2015, 2015, pp. 58-60.
10. Kato, J., Ogata, M., Inoue, T., Goto, M.. Songle Sync: A Large-Scale Web-based Platform for Controlling Various Devices in Synchronization with Music, Proc. of MM 2018, 2018, pp. 1697-1705.
11. Goto, M., Yoshii, K., Fujihara, H., Mauch, M., Nakano, T.. Songle: A Web Service for Active Music Listening Improved by User Contributions, Proc. of ISMIR 2011, 2011, pp. 311-316.
12. 小幡朱, 串山久美子. ステージパフォーマンスにおいて観客と演者の コミュニケーションを支援する LED を使った衣装の提案, インタラクション2017, 2017.
13. 川元留輝, 串山久美子. コンサートで観客の行動を促す演出と LED デバイス, インタラクション2018, 2018.
14. 高橋佑汰, 松本康介, 鈴木莉野, 相良亮, 河村逸平, 内藤唯香, 田口愛祐美, 中村桃子, 落合美礼, 上松大輝. 理想のライブを現実に！～観客と演者の両方が演出する空間, インタラクション2020, 2020.
15. 岩本裕磨, 岩井将行. 音楽イベント初心者の応援行動の同期性の向上する無線通信機能搭載型ペンライト, エンタテインメントコンピューティングシンポジウム2016論文集, 2016, pp. 311-312.
16. 武井秀憲, 山田篤志, 越後宏紀, 菅野真功, 小林稔. 音楽イベントの応援行動を誘導するペンライト型デバイスの発光パターンの検討と体験システム, インタラクション2020, 2020.
17. 平川新吾, 末光智彦, 小林威晴, 佐藤皇太郎, 大久保明. ライブ・エンタテインメントコミュニケーションシステム「LumiConne（ルミコネ）」の開発経緯と実装技術, 研究報告エンタテインメントコンピューティング（EC）, 2012, vol. 2012-EC-23, no. 4, pp. 1-6.
18. 堂林まどか, 沖真帆, 塚田浩二. Sync ☆ Idol：ライブアイドルとファンを盛り上げるライブ支援装置, WISS 2015, 2015.
19. Kayali, F., Hödl,O., Fitzpatrick, G., Purgathofer, P., Filipp, A., Mateus-Berr, R., Kühn, U., Wagensommerer, T., Kretz, K., Kirchmayr, S.. Playful Technology-Mediated Audience Participation in a Live Music Event, Pub. of CHI PLAY EA 2017, 2017, pp. 437-443.
20. amazarashi. https://www.amazarashi.com/, (参照 2021-01-23).
21. THE ORAL CIGARETTES. https://theoralcigarettes.com/, (参照 2021-01-23).
22. 音楽ナタリー, 「amazarashi、日本武道館公演で繰り広げた圧巻の“朗読演奏実験空間”」,  
    https://natalie.mu/music/news/308415, (参照 2021-01-23).
23. 音楽ナタリー, 「オーラル、今伝えたいことを表現した有観客公演「生のライブに代わるものはない」」,  
    https://natalie.mu/music/news/397986, (参照 2021-01-23).
24. 紫村勇綺, 宮脇健三郎. ジェスチャ認識による演奏者と観客のインタラクションに基づくライブの一体感演出システム, インタラクション2019, 2019.
25. Cao, Z., Hidalgo, G., Simon, T., Wei, S., Sheikh, Y.. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, Proc. of CVPR 2017, 2017.
26. 寺内涼太, 福島史康, 大津耕陽, 福田悠人, 小林貴訓, 久野義徳, 山崎敬一. 一体感を増強する遠隔ライブ参加システム, インタラクション2019, 2019.
27. 米澤拓郎 , 徳田英幸. Control Manually: 視聴者協力型ライブ演出システムによるコミュニケーションと演出効果の拡張, 研究報告エンタテインメントコンピューティング（EC）, 2012, vol. 2012-EC-25, no. 8, pp. 1-8.
28. 浜中雅俊, 李昇姫. コンサートスコープヘッドホン, 日本バーチャルリアリティ学会論文誌, 2013, vol. 18, no. 3, pp. 227-236.
29. 吉田有花, 宮下芳明. 身体動作の重畳表示による動画上での一体感共有, インタラクション2012, 2012.
30. 粕谷貴司, 塚田学, 菰原裕, 高坂茂樹, 水野拓宏, 野村譲誉, 上田雄太, 江崎浩. インタラクティブな遠隔ライブVR配信プラットフォーム, 情報処理学会論文誌デジタルコンテンツ（DCON）, 2019, vol. 7, no. 2, pp. 1-14.
31. Kaneko, T., Tarumi, H., Kataoka, K., Kobuchi, Y., Yamashita, D., Nakai, T., Yamaguchi, R.. Supporting the Sense of Unity between Remote Audiences in VR-Based Remote Live Music Support System KSA2, Proc. of AIVR 2018, 2018, pp. 124-127.
32. Hudson, S.. Metal Movements: Headbanging as a Legacy of African American Dance, Modern heavy metal : markets, practices and cultures : International Academic Research Conference, 2015.
33. Walser, R.. Running With The Devil: Power, Gender and Madness in Heavy Metal, Wesleyan University Press, 1993.
34. Metallica. https://www.metallica.com/, (参照 2021-01-23).
35. Pillsbury, G.. Damage Incorporated Metallica and the Production of Musical Identity, New York: Routledge, 2006.
36. Purcell, N. J.. Death Metal Music: The Passion and Politics of a Subculture, Jefferson, North Carolina: McFarland & Company, 2003.
37. Bardos, L., Korinek, S., Lee, E., Borchers, J.. Bangarama: Creating Music With Headbanging, Proc. of NIME 2005, 2005, pp. 180-183.
38. Merrill, D.. Head-tracking for gestural and continuous control of parameterized audio effects, Proc. of NIME 2003, 2003, pp. 218-219.
39. Moll, P., Leibetseder, A., Kletz, S., Lux, M., Muenzer, B.. Alternative inputs for games and AR/VR applications: deep headbanging on the web, Proc. of MMSys 2019, 2019, pp. 320-323.
40. PoseNet.   
    https://github.com/tensorflow/tfjs-models/tree/master/posenet, (参照 2021-01-23).
41. Papandreou, G., Zhu, T., Kanazawa, N., Toshev, A., Tompson, J., Bregler, C., Murphy, K.. Towards Accurate Multi-person Pose Estimation in the Wild, Proc. of CVPR 2017, 2017, pp. 3711-3719.
42. Pei, L., Guinness, R., Chen, R., Liu, J., Kuusniemi, H., Chen, Y., Chen, L., Kaistinen, J.. Human Behavior Cognition Using Smartphone Sensors, Sensors, 2014, vol. 13, no. 2, pp. 1402-1424.
43. Anjum, A., Ilyas, M. U.. Activity recognition using smartphone sensors, Proc. of CCNC 2013, 2013, pp. 914-919.
44. Lee, H., Choi, Y. S., Lee, S., Shim, E.. Smart pose: mobile posture-aware system for lowering physical health risk of smartphone users, Proc. of CHI EA 2013, 2013, pp. 2257-2266.
45. Nipun, D. N., Reza, A., Amir, H. B.. Ergonomic analysis of construction worker's body postures using wearable mobile sensors, Applied Ergonomics, 2017, vol. 62, pp. 107-117.
46. 安永卓哉, 中澤篤志, 竹村治雄. 加速度センサによるユーザコントロールを導入した音楽に合った舞踊動作の自動生成, 研究報告音声言語情報処理（SLP）, 2012, vol. 2012-SLP-90, no. 25, pp. 1-6.
47. Kanke, H., Takegawa, Y., Terada, T., Tsukamoto, M.. Airstic Drum: A Drumstick for Integration of Real and Virtual Drums, Proc. of ACE 2012, 2012, pp. 57-69.
48. Ladha, C., Hammerla, N., Hughes, E., Olivier, P., Ploetz, T.. Dog’s Life: Wearable Activity Recognition for Dogs, Proc. of UbiComp 2013, 2013, pp. 415-418.
49. Fear, and Loathing in Las Vegas. http://www.lasvegas-jp.com/, (参照 2021-01-23).
50. Crossfaith. http://crossfaith.jp/, (参照 2021-01-23).
51. coldrain. http://coldrain.jp/, (参照 2021-01-23).
52. SiM. https://sxixm.com/, (参照 2021-01-23).
53. Survive Said The Prophet. https://survivesaidtheprophet.com/, (参照 2021-01-23).
54. scikit-learn. https://scikit-learn.org/, (参照 2021-01-23).

**著者紹介**

 **二宮洸太**

明治大学総合数理学部先端メディアサイエンス学科在学．手書き文字の可読化やコミック，ヘドバンに関する研究に着手．多趣味．