

楽器演奏における身体動作と音データの同時計測と解析手法の開発

研究代表者 柴田 傑 室蘭工業大学 しくみ解明系領域 助教
研究協力者 角 美弥子 北海道教育大学 芸術・スポーツ文化学科 准教授

1 はじめに

楽器演奏をはじめとする身体技能は、学習者が技能における重要な動きを“コツ”として獲得し、繰り返して訓練することによって、学習者の身体に合うように最適化して身に付けると考えられる。しかしながら、民俗芸能や和楽器の演奏における基本的な動作は、学習方法の体系化、コツの言語化がされていないものも多い。そのため、一般に「見て学ぶ」「技を盗む」といわれるように、学習者は観察と模倣およびそれらを確認する反復練習によって自ら動作のコツをつかみ、動きを習得することとなる。

本研究では、観察と模倣によって学習を進める環境において、情報・電気通信技術を用いてコツを解析する手法について検討し、学習者のより早い技能習得を可能にする補助的な情報を提示することを目指す。特に日本の芸能において重要な役割を果たす和楽器である三味線に着目する。三味線演奏の身体動作と、演奏の評価対象となる音を同時に計測し、言語化が難しいコツを解析する。

筆者らは民俗芸能の舞踊を対象としモーションキャプチャ (MoCap) を活用した学習支援システム[1, 2]、バーチャルリアリティ (VR) やプロジェクションマッピングを用いた三味線の学習支援システムを構築してきた[3, 4]。これらの手法は、学習者に対し、CG を用いて熟練者の模範的な動きを提示することによって、身体技能の模倣を支援する。模範的な動きの提示は、動きの全体を把握できていない学習の最初期段階において一定の効果が得られることがわかった。しかしながら、学習を進めるためには、動きの外見的な模倣を通して、学習者が動きの特徴を理解し、動きを客観的にとらえて修正できることが重要になると考えられる。特に楽器演奏は身体技能であるものの、演奏によって発せられる音が評価の対象となる。そのため、身体動作と音の関連性を意識した学習が不可欠である。

現在、三味線の演奏を支援する研究として、演奏を保存するための自動採譜システム[5]や演奏に演奏者の負担を減らすための譜面の自動めくりシステム[6]が提案されている。これらの研究は主に発せられる音に着目した研究である。本研究では、1) 複数のセンサを用いて身体の動き、力の入れ具合、および音を同時計測するシステムの開発、2) 計測した複数センサのデータを解析することによって身体動作と音の関係についての知見を得ることを目的とする。

本研究の成果は、熟練者のコツを学習者に示すことによって模倣を支援する、学習者の動きを解析してフィードバックし訓練を支援するなど、電気通信技術を用いた身体技能の学習支援への応用が期待できる。

2 身体動作と音の同時計測システム

2-1 同時計測システム概要

三味線は、リュート属の撥弦楽器であり、胴と棹が接続された形状をしている。胴は空洞に皮が張られた太鼓の構造をしており、胴から棹にかけて3本の弦が張られている。演奏時には、胴を右太ももに置き、左手の指先または爪先で弦を押さえて音程を決定し、右手の撥で胴と弦を同時に叩いて演奏する。三味線の演奏では、三味線と撥の相対的な位置・姿勢について、右腕で制御することによってさまざまな音を発する。そのため、三味線の学習では、右腕の速度・加速度、力の入れ具合といった右腕の動かし方が重要になると考えられる[7]。

現在ギターやバイオリンを対象に複数センサを用いてデータを計測する研究が進められているものの[8, 9]、熟練者のデータのアーカイブを目的としており、学習者が練習目的で利用することは想定されていない。本研究では、学習の支援を想定し1台の制御用のコンピュータで操作可能なシステムを開発し、熟練者および初心者の演奏を計測する。開発システムは、磁気式 MoCap を用いて三味線と撥の位置・姿勢 (6 DOF)、慣性センサを用いた右腕の加速度・角速度 (6 DOF)、筋電位センサを用いて右腕の筋活動量 (8 CH)、ピックアップマイクを用いて三味線が発する音 (1 CH) を計測する。

2-2 ハードウェア構成

図1に提案する同時計測システムを示す。本システムは、実物の三味線、撥、磁気式 MoCap (Polhemus 社 Liberty), ジェスチャ入力デバイス (GestureCtr: ThalmicMyo 社 Myo Armband), ピックアップマイク (Mic: SHADOW 社 SH711), 制御用のコンピュータ (制御用 PC: Dell Precision M3800) で構成される。

磁気式 MoCap (同図 (a)) はトランスミッタ (同図 (b)) が発する磁気をセンサに内蔵されたコイルで検知することによって、トランスミッタを原点とする各センサの位置 (3 DOF), 姿勢 (3 DOF) を計測するシステムである。MoCap のセンサを同図 (c) に示す撥の中央, および同図 (d) に示す三味線の胴の横の二カ所に装着し、三味線と撥の位置・姿勢を計測する。磁気式 MoCap は USB ケーブルによって制御用 PC に接続する。

GestureCtr (同図 (e)) は、表面筋電位センサ (EMG: 8 CH) および慣性センサ (IMU: 加速度 3 DOF, 角速度 3 DOF) を内蔵したコントローラデバイスであり、演奏者の力の入れ具合を計測する。GestureCtr は演奏時に撥を持つ右腕の前腕に装着する。GestureCtr は Bluetooth 通信によって制御用 PC に接続する。

Mic は楽器に張り付け可能な薄型のマイクであり、胴の振動 (VOL: 1 CH) を録音する。Mic は同図 (f) に示す胴の後ろ部分に装着する。Mic は USB ケーブルによって制御用 PC に接続する。MoCap のセンサおよび Mic は三味線熟練者の指導に従い、三味線を構えたときに干渉しない位置を決定した。装着にはテープおよび 3D プリンタで作成した治具を用いた。

本システムは 1 回で合計 27 次元のデータ (1 フレーム) を取得でき、1 秒間に約 60 回 (Mic は 44.1 kHz) データを保存する。

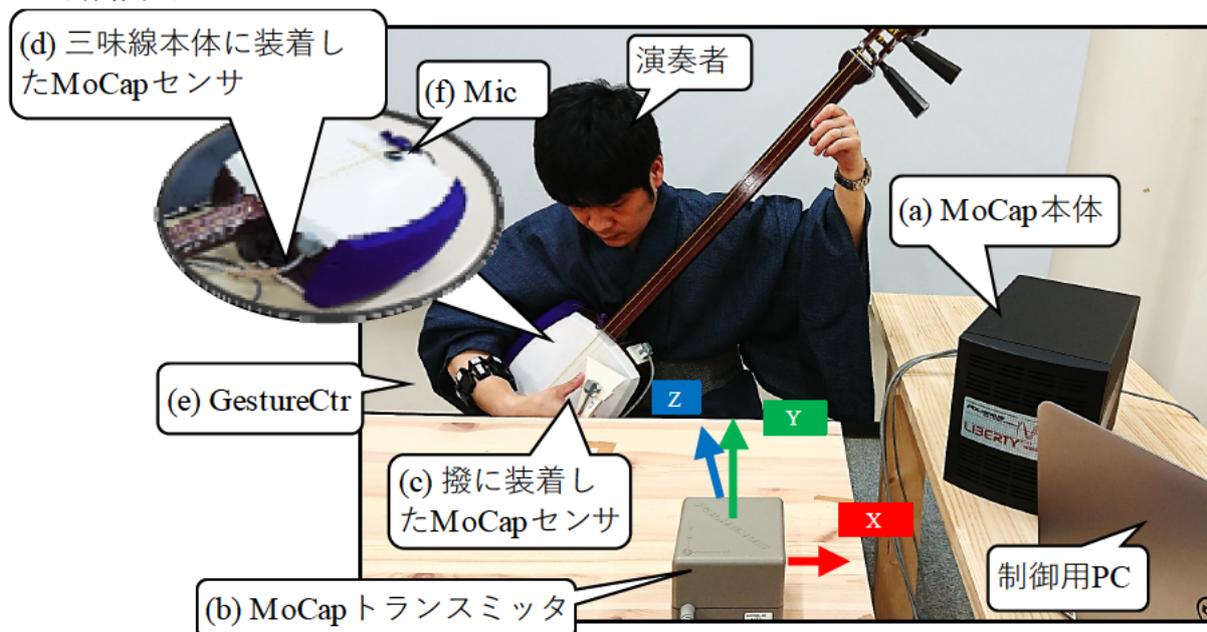


図1 ハードウェア構成

2-2 ソフトウェア構成

本システムのソフトウェアは、VR アプリケーションの開発環境である Unity で開発した。本ソフトウェアはリプレイモード、リアルタイムモード、計測モードの3つのモードを提供する。本ソフトウェアの各モードおよびリプレイモードの実行時の画面を図2に示す。同図に示すように、キーボードの「S キー」, 「E キー」および「P キー」の押下によって、それぞれのモードへ移行する。

リプレイモードでは、画面上部の LOAD ボタンから計測済のデータを読み込み、動き、力の入れ具合、音を確認することができる。画面上部には三味線、撥、基準となるトランスミッタの CG が表示される。また、マウス操作を用いて、視点を変えて観察することができる。画面上部のシークバーを操作するとデータの再生時刻を自由に変更できる。画面下部には、レーダーチャートが表示され、IMU, および EMG の値が表示される。リプレイモードでキーボードの「P キー」が押下されるとリアルタイムモードへ移行する。

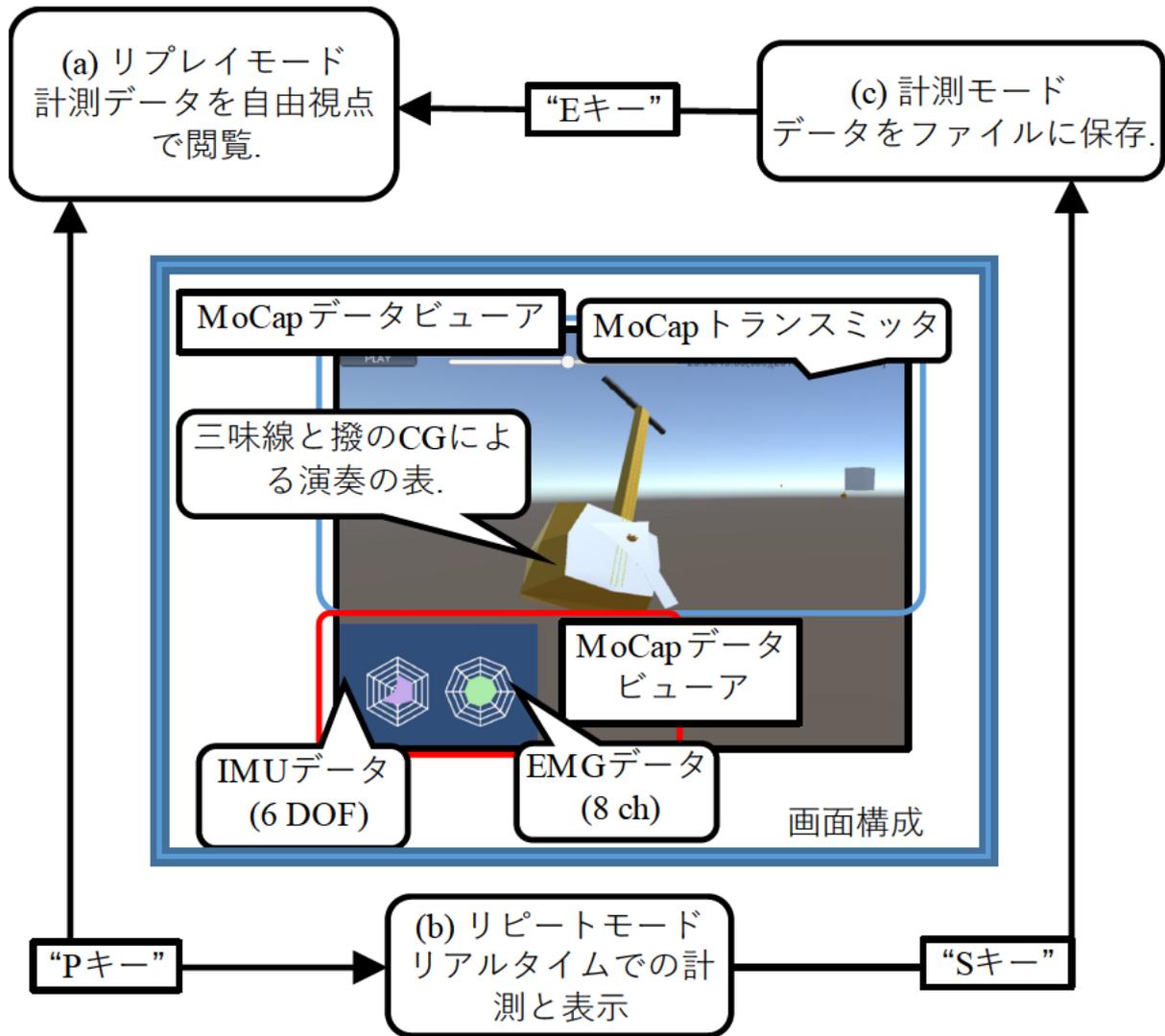


図2 ソフトウェア構成

リアルタイムモードは、リプレイモードと同じCGとレーダーチャートを用いて、リアルタイムに三味線と撥の動きと力の入れ具合を表示する。表示された動きは、マウス操作を用いて、視点を変えて観察することができる。リアルタイムモードで、キーボードの「Pキー」が押下されるとリプレイモード、キーボードの「Sキー」が押下されると、計測モードへ移行する。

計測モードでは、リアルタイムモードの表示を維持したまま、接続されたすべてのセンサが同時に計測を開始する。MoCapおよびGestureCtrはCSV形式、Micは44.1kHzのWAV形式で計測したデータをファイルに保存する。このとき、ファイル名は計測の開始日時となる。計測モード中にキーボードの「Eキー」が押下されると計測を終了し、リプレイモードへ移行する。

2-3 計測実験

(1) 計測実験環境

実験では、熟練者1名(津軽三味線:経験8年/指導の資格あり)、経験者1名(地歌:経験2年/2月に一度の受講)、初心者4名(大学生/初めて三味線に触れる)の演奏を計測した。演奏者は、まず、右腕にGestureCtrを装着し、キャリブレーション用のジェスチャを入力する。キャリブレーションは、GestureCtrに付属のソフトウェアを用いた。その後、MoCapセンサが装着された三味線と撥を構え、演奏を開始する。計測時の座標の基準となるトランスミッタは演奏者の正面の机に設置した。このとき、MoCapの各軸は、演奏者の左方向をX軸、上方向をY軸、後ろ方向をZ軸とした。熟練者のデータを計測する最初の実験では、提案システ

ムの実行時の画面を熟練者の正面にプロジェクタで投影した。これによって、熟練者が各モードの様子を確認しながら演奏できるようにした。また、実験の様子は、記録用のビデオカメラで撮影した。

実験では、初めて三味線に触る初心者は楽曲の演奏ができないので、撥さばきの基本的な動作として、90 bpm のテンポで1つの音を鳴らす図3に示すようなパターンを計測した。同図のパターンを4回繰り返す演奏を1フレーズとして、3から4フレーズを1回の計測(1 take)とした。参加者の状況に応じ、それぞれ3から5 take計測した。また、初心者は実験前に、三味線の教則本[10]を参考に三味線の構え方と撥の動かし方を10分程度学ぶこととした。熟練者の演奏では、基本的な動作に加えて、初心者向けの楽曲である“さくらさくら”，および多彩な撥さばきが含まれる上級者向けの楽曲“津軽じょんがら節”をそれぞれ3 take計測した。



図3 演奏パターン

計測の結果、MoCap および GestureCtr は平均 57.6 Hz (標準偏差 6.77) のサンプリングレートで計測できた。計測後に提案システムのリプレイモードを用いて、計測したデータを再生してビデオカメラによる動画像と目視で比較したところ、動きと音に大きなずれは感じられなかった。

(2) 計測実験結果

図4に“さくらさくら”の計測結果の一部を示す。横軸は時間(フレーム)、縦軸は正規化されたデータの値である。同図の点線の折れ線グラフはMoCapで計測した撥の高さ、同図の実線の折れ線グラフはIMUで計測した右腕の上下方向の加速度、同図の破線の折れ線グラフは音量の二乗平均平方の対数である。音量は二乗平均平方の区間をMoCapの計測時刻の近傍とすることによって、リサンプリングしている。同図にはタイミンングの目安として歌詞を示した。

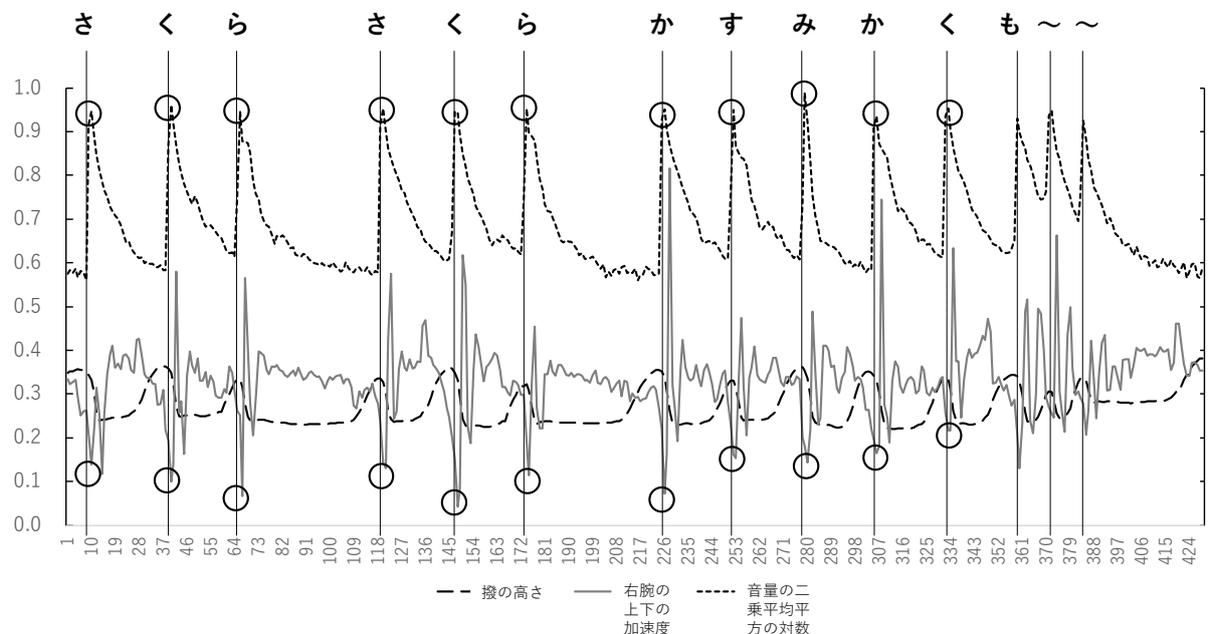


図4 “さくらさくら”の計測結果

同図の丸で囲った部分に示すように、音量の正のピークに近い時刻で右腕の加速度の負のピークが観測できる。これは、打ち下ろす方向に力が加わった直後に音が鳴っていることを意味する。また、撥の高さと比

較すると、撥が一定の高さから下に移動し、完全に振り下ろされるまでの間で弦に触れて音が鳴っていることがわかる。

“さくらさくら”の take 1 について、比較的腕の振りが大きい四分音符が連続するパターン (36 カ所) において、音量と右腕の加速度のピークを手動で抽出し、時間差を求めた。時間差の平均は 0.95 フレームであり、最もズレが大きい箇所は二カ所で、4 フレームのズレがあった。音量と右腕の加速度のピークが完全に一致していない理由として、二乗平均平方を用いてサンプリングレートを合わせる処理を施していること、計測用 PC の CG 表示などに伴う計算負荷などによってサンプリング間隔にバラつきがあること、撥の動きは腕だけでなく手首の動きとの連携で実現していることなどが考えられる。しかしながら、時間差の平均は 1 フレーム以内であり、解析に用いる特徴が計測できていると考える。

以上の結果から、MoCap, GestureCtr, Mic で音になるタイミングが一致しており、同期してデータを計測できていると考えられる。

3 身体動作と音の関連性の解析

3-1 解析手法概要

三味線の学習においては、目標となる演奏と自身の演奏との違いを客観的に把握、修正することが重要と考えられる。三味線の学習を支援するために、開発した同時計測システムを用いて計測した三味線演奏の解析手法を検討し、演奏の習得に寄与する特徴の抽出手法を検討する。本研究では、計測したデータから演奏の違いを抽出する手法としてニューラルネットワーク (NN) に着目する。NN は与えられたデータから分類に寄与する特徴量と分類基準を自動的に抽出する手法である。まず、予備的な実験として同時計測したデータの一部を抽出して組み合わせ、NN によって分類することによって、抽出されたデータが分類にどのように寄与しているか検討する。次に、計測したセンサのデータ間の関連性を考慮した分類手法について検討する。

3-2 計測データの特徴に関する予備実験

(1) 特徴抽出

本研究では、ノイズやバラつきのある画像データを効果的に分類可能な畳み込みニューラルネットワーク (CNN) に着目する。CNN は神経細胞をモデル化した計算による分類手法で、重み付きの経路が接続された構造に従ってデータを処理することによって分類する手法である。CNN では、あらかじめ分類結果が与えられたデータセット (学習データ) を用いて、正しい分類結果が得られるように経路の重みを決定 (学習) する。学習が完了した CNN に未知のデータ (テストデータ) を入力することによって分類する。CNN は同一サイズの画像を想定した NN であり、入力するデータは行数と列数が同一の二次元データである必要がある。

一般に MoCap データは時系列データであり、計測時間によってデータサイズは変動する。これに対し、計測した MoCap データを同一サイズになるように分割し、分割されたデータを用いることによって身体動作を分類する試みが進められている [11]。本システムを用いて計測したデータは、計測時間が n フレームであるとき、 n 行 27 列の行列として記録される。そこで、予備的な実験として、計測したデータを分割し、一部を抽出してから同一サイズになるように分類し、分類結果を比較する。これによって、抽出したデータに分類に寄与するデータが含まれているか明らかにする。

(2) 予備実験 1: 楽曲の違いに関する予備実験

予備実験 1 では、実験で計測した熟練者の演奏について、“さくらさくら”と“津軽じょんがら節”で身体の動かし方の違いが観測できているかどうか明らかにする。そこで計測データを“さくらさくら”と“津軽じょんがら節”の 2 クラスに分類する。まず、熟練者が演奏したデータを同図 (a) に示すように、15 フレーム (約 0.25 秒) 毎に分割した。分割した結果 327 個のデータが得られた。計測したデータの各列は、計測したセンサのチャンネルごとに、図 5 (b) に示す 8 種類のデータとみなす。8 種類から 3 つ選んだ組み合わせおよびすべてのセンサ (計 57 種類) について CNN で分類して比較する。CNN の学習では、1 take 分から抽出し、分割したデータを学習データとし、残りの 2 take から抽出されたデータをテストデータとした。CNN の各パラメータは表 1 のとおりである。

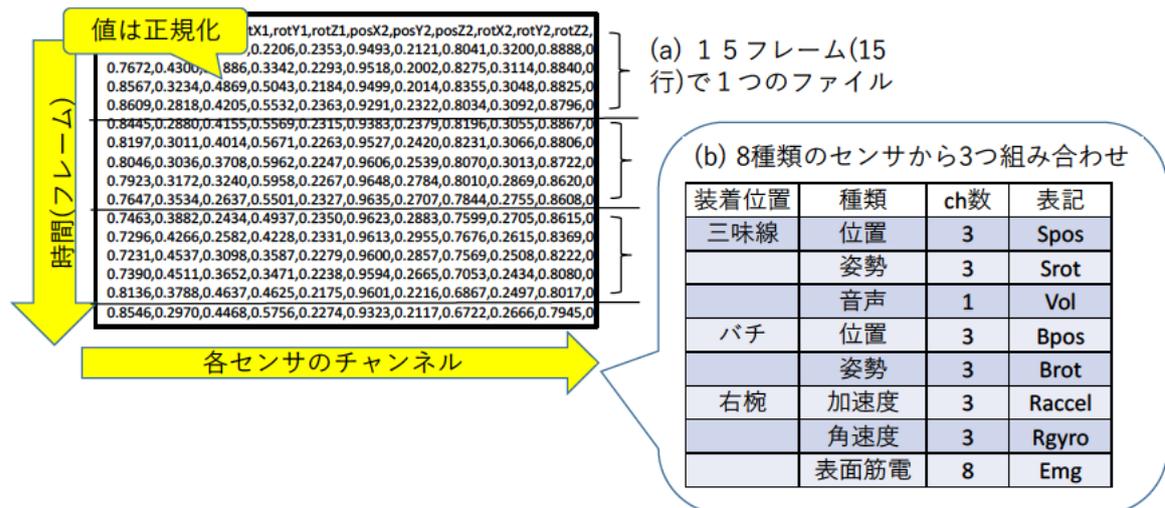


図5 データの分割

表1 CNNのパラメータ

学習回数	10000	
学習率	0.0001	
活性化関数	ReLU関数	
勾配法	Adam	
ミニバッチサイズ	10	
畳み込み層	第一層(Conv1)	第二層(Conv2)
フィルター数	32	64
カーネルサイズ	5×5	5×5
ストライド	1	1
プーリング層		
プーリング手法	マックスプーリング	
カーネルサイズ	2×2	
ストライド	2	

分類結果は分類における再現率と適合率の調和平均である f 値で評価する。図6に得られた f 値を示す。縦軸は f 値、横軸はセンサの組み合わせを表し、“さくらさくら”の f 値と“津軽じょんがら節”の f 値の和を降順に並べた。同図から全ての組み合わせで、“さくらさくら”よりも“津軽じょんがら節”の f 値が高いことがわかる。最も f 値の高い組み合わせ (MAX) は Bpos (撥の位置), Srot (三味線の角度), Rgyro (腕の角速度) であった。すべてのセンサを使った場合は、57種類中42番目であった。以上の結果から、撥の位置および三味線の角度には“さくらさくら”と“津軽じょんがら節”の違いはMAXの組み合わせによく表れていると考えられる。

(3) 予備実験2：演奏経験毎の違いに関する予備実験

予備実験2では計測した熟練者、経験者、初心者の基本的な動作の演奏について、演奏の経験者毎のクラスにCNNを用いて分類できるか確認する。このとき、データの時間分割の方法を変えることによって、時間的な特徴について検討する。本実験で計測した動作は一定のテンポで繰り返すことを目標とする動作である。そのため、データ分割・抽出の前処理において、繰り返しの特徴が保たれるように分割する必要がある。しかしながら、初心者の動きは安定していないので、効果的に分類可能な分割時間(適切フレーム数)は明らかではない。そこで本実験では、分割するフレーム数を変えてCNNで分類し、分割の仕方が分類に与える影響について検討した。

本研究では、初心者59フレーズ、経験者19フレーズ、上級者16フレーズを対象とした。実験では、セ

ンサに対する前処理として、文献[3]に基づいて、MoCap センサを用いて、地面对して棹の方向がなす角度（棹の角度：sAng）、胴の法線がなす角度（胴の角度：dAng）、三味線に対する撥の相対的な位置と姿勢を求め、EMG、IMU、音量を加えた 23 次元のデータとした。1 フレーズを分割するサイズ（分割フレーム数）を N とし、分割フレーム数 N で分割した $N \times 23$ 行列のデータ（分割データ）を使い、CNN によって分類した。実験に使う分割フレーム数は、15 から 60 まで 5 刻みに増やし、合計 10 パターンで比較した。CNN は楽曲の違いに関する予備実験と同じ表 1 に示すパラメータを用いた。評価は、クラス毎に求めた f 値の相加平均とした。

求めた f 値を図 6 に示す。同図の横軸は分割フレーム数、縦軸は f 値の平均である。実験の結果、分割フレーム数 25 が最も f 値の平均が高かった。しかしながら、最も低い f 値であっても 0.91 と一定の精度であることが分かった。

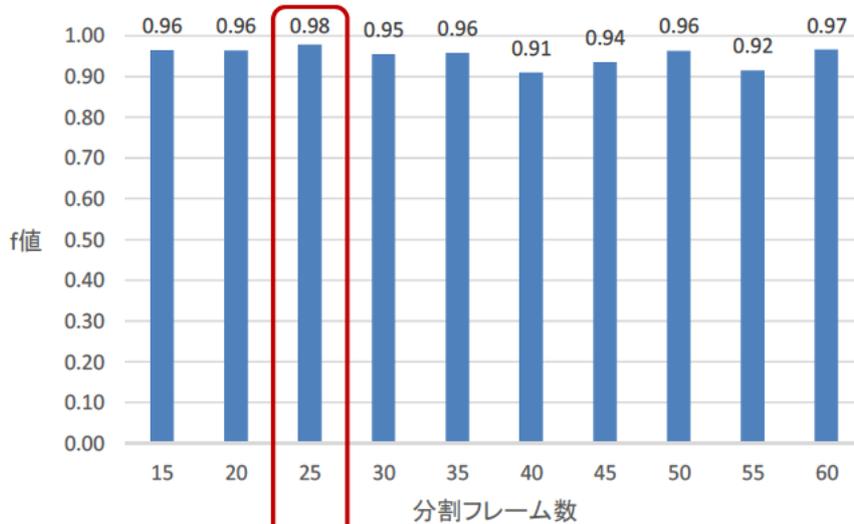


図 6 演奏経験毎の分類結果

2-2 身体の使い方と音データの解析

(1) 前処理

本解析では、詳細な解析を進めるため、基本的な動作 1 回毎について演奏者の違いの分類を試みる。計測したデータについて、音を一回出している区間を手動で抽出した。演奏の失敗や計測のノイズを除き、85 回分のデータが得られた。本実験では、身体動作と音の関係を解析するために、音声データはバンドパスフィルタで 2 オクターブずつ 5 つのバンドに分解した。得られたデータは、演奏時の演奏者の状態の時系列である。そこで、時間変化に対する“身体の使い方”を比較するために、データの相関行列を求めた。相関行列によって、演奏者がどの部位を活用して演奏しているのかを比較できると考える。また、相関行列はサイズが同一の二次元データとして扱うことができるので、CNN での分類にも適していると考えられる。

本実験で用いた前処理は以下の通りである。

【MoCap のデータ】

1. 三味線の棹の角度・胴の角度 (sAng, dAng) を求める。
2. 三味線に対する撥の位置姿勢 (bPosX, bPosY, bPosZ, bRotX, bRotY, bRotZ) を求める。

【GestureCtr のデータ】

1. EMG の絶対値をとって整流化する (emg. 1~emg. 8)。
2. EMG の総合的な力の入れ具合を評価するために、8 CH のノルムを求める (emgAll)
3. IMU は計測データを用いる (accelX, accelY, accelZ, gyroX, gyroY, gyroZ)。

【Mic データ】

1. FIR バンドパスフィルタを用いて 440Hz を基準に 2 オクターブ毎の周波数帯に分離する。
2. MoCap のサンプリング時刻を基準に二乗平方平均の log をとってリサンプリングする (band.all, band. 1~band. 5)。

[データ間の関係の抽出]

1. 相関行列を求める.

(2) 相関行列による比較

表2に求めた相関行列の一部を示す. 同表の塗りつぶされたセルは相関係数の絶対値が0.7以上, 色付き文字は相関係数の絶対値が0.4以上であることを示している. 同表から, 熟練者は腕の角速度 (gyroX) と音声との相関が高いのに対し, 初心者は撥の位置 (bPosX, Y, Z) と音声の相関が高い傾向がある. これは, 熟練者が腕の回転によって撥を動かしているのに対し, 初心者は撥の位置のみによって演奏していることを示唆している. 熟練者は, 筋電の CH 1 (emg.1) と音声との相関も高いが, 初心者は各筋電の間での相関が高いものの, 音声との相関は高くない. これは, 初心者は不要な部分に力が入り, 撥に力が伝えられていない可能性があることを示している.

表2 相関係数の比較結果

(a) 熟練者の相関行列

三味線の状態	撥の位置/姿勢							筋電位								加速度/角速度				音声									
	sAng	dAng	bPosX	bPosY	bPosZ	bRotX	bRotZ	emg.1	emg.2	emg.3	emg.4	emg.5	emg.6	emg.7	emg.8	emgALL	acceX	acceY	acceZ	gyroX	gyroY	gyroZ	lendALL	band.0	band.1	band.2	band.3	band.	
sAng	1.00	-0.75	-0.63	0.70	0.53	0.18	0.2	0.48	-0.17	0.19	0.32	0.12	-0.01	0.13	0.0	-0.25	0.17	0.06	-0.07	-0.22	-0.15	0.05	-0.12	-0.07	-0.03	-0.11	-0.07	-0.09	0.02
dAng	-0.75	1.00	0.52	-0.62	-0.45	-0.22	-0.30	-0.47	-0.09	-0.23	-0.33	-0.11	-0.22	0.0	-0.21	0.02	-0.27	-0.07	0.11	0.1	-0.1	0.23	-0.0	-0.30	-0.16	-0.31	-0.30	-0.31	-0.27
bPosX	-0.63	0.52	1.00	-0.97	-0.97	-0.1	-0.63	-0.72	0.15	-0.07	-0.2	-0.05	0.12	-0.1	-0.03	0.21	-0.09	-0.10	0.19	0.55	0.22	0.30	0.33	0.23	-0.05	0.02	0.23	0.29	0.00
bPosY	0.70	-0.62	-0.97	1.00	0.95	0.3	0.62	0.74	-0.11	0.12	0.26	0.09	-0.10	0.13	0.05	-0.18	0.12	0.13	-0.22	-0.96	-0.12	-0.2	-0.25	-0.10	0.11	0.09	-0.10	-0.17	0.10
bPosZ	0.53	-0.45	-0.97	0.95	1.00	0.39	0.71	0.72	-0.09	0.06	0.2	0.0	-0.12	0.12	0.0	-0.1	0.09	0.1	-0.27	-0.97	-0.06	-0.33	-0.25	-0.1	0.10	0.01	-0.1	-0.20	0.07
bRotX	0.18	-0.22	-0.3	0.3	0.39	1.00	0.62	0.19	-0.0	0.12	0.27	-0.01	-0.05	-0.09	-0.0	-0.09	0.07	0.01	-0.02	-0.17	0.10	-0.25	-0.06	-0.21	-0.0	-0.18	-0.21	-0.23	-0.16
bRotY	0.2	-0.30	-0.63	0.62	0.71	0.62	1.00	0.55	0.0	0.10	0.26	-0.10	0.07	0.12	0.17	0.03	0.11	0.07	-0.15	-0.35	0.18	-0.50	-0.13	-0.10	0.25	0.07	-0.10	-0.1	0.11
bRotZ	0.48	-0.47	-0.72	0.74	0.72	0.19	0.55	1.00	0.16	0.08	0.17	0.06	0.06	0.15	0.18	0.01	0.17	-0.08	0.0	-0.20	0.15	-0.33	-0.01	0.03	0.16	0.15	0.0	-0.02	0.17
emg.1	0.17	-0.09	0.15	-0.11	-0.09	-0.0	0.0	0.16	1.00	0.22	0.0	-0.10	0.13	-0.02	0.11	0.52	0.18	-0.06	0.13	0.20	0.43	-0.18	0.41	0.46	0.43	0.42	0.46	0.47	0.43
emg.2	0.19	-0.23	-0.07	0.12	0.06	0.12	0.10	0.22	0.22	1.00	0.50	0.28	0.08	-0.02	0.07	-0.06	0.51	0.00	-0.10	0.08	0.08	-0.06	0.01	0.22	0.25	0.2	0.22	0.16	0.28
emg.3	0.32	-0.33	-0.2	0.26	0.2	0.27	0.26	0.17	0.0	0.50	1.00	0.17	0.03	-0.01	0.06	-0.06	0.60	0.10	-0.1	-0.09	0.11	-0.10	0.06	0.10	-0.1	0.10	0.10	0.08	0.15
emg.4	0.12	-0.11	-0.05	0.09	0.0	-0.01	-0.10	0.06	-0.10	0.28	0.17	1.00	0.1	0.11	-0.11	-0.13	0.70	0.15	-0.1	-0.0	0.02	-0.13	-0.07	0.10	0.05	0.11	0.09	0.11	
emg.5	-0.01	-0.22	0.12	-0.10	-0.12	-0.05	0.07	0.06	0.13	0.08	0.03	0.1	1.00	-0.02	0.23	0.3	0.38	-0.10	0.06	0.31	0.25	-0.30	0.00	0.31	0.2	0.3	0.30	0.3	0.2
emg.6	0.13	-0.0	-0.1	0.13	0.12	-0.09	0.12	0.15	-0.01	-0.02	-0.01	0.11	-0.02	1.00	0.23	-0.02	0.29	-0.02	0.06	-0.01	-0.05	-0.07	-0.10	-0.07	0.05	-0.01	-0.07	-0.08	0.01
emg.7	0.0	-0.21	-0.03	0.05	0.0	-0.0	0.17	0.18	0.32	0.07	0.06	-0.11	0.23	0.23	1.00	0.17	0.17	-0.09	0.17	0.12	0.21	-0.10	0.20	0.22	0.11	0.26	0.21	0.22	0.17
emg.8	0.25	0.02	0.21	-0.18	-0.11	-0.09	0.03	0.01	0.52	-0.02	-0.06	-0.13	0.13	-0.02	0.17	1.00	0.25	0.06	0.09	0.25	0.44	-0.29	0.27	0.3	0.31	0.35	0.3	0.37	0.30
emgALL	0.17	-0.27	-0.09	0.12	0.09	0.07	0.11	0.17	0.18	0.51	0.60	0.70	0.38	0.29	0.17	0.25	1.00	0.09	-0.06	0.10	0.21	-0.27	0.0	0.2	0.23	0.28	0.2	0.2	0.27
acceX	0.06	-0.07	-0.10	0.13	0.1	0.01	0.07	-0.08	-0.06	0.00	0.10	0.15	-0.10	-0.02	-0.09	0.06	0.09	1.00	-0.60	-0.0	0.08	-0.25	-0.09	0.15	0.17	0.12	0.16	0.13	0.18
acceY	-0.07	0.11	0.19	-0.22	-0.27	-0.02	-0.15	0.0	0.13	-0.10	-0.1	-0.1	0.06	0.06	0.17	0.09	-0.06	-0.60	1.00	0.48	-0.11	0.22	0.2	-0.27	-0.26	-0.12	-0.27	-0.23	-0.31
acceZ	-0.22	0.1	0.55	-0.56	-0.57	-0.17	-0.35	-0.20	0.20	0.08	-0.09	-0.0	0.31	-0.01	0.12	0.25	0.10	-0.0	0.48	1.00	0.15	0.05	0.18	0.17	-0.02	0.0	0.16	0.22	0.03
gyroX	-0.15	-0.1	0.22	-0.12	-0.06	0.10	0.18	0.15	0.43	0.08	0.11	0.02	0.25	-0.05	0.21	0.41	0.21	0.98	-0.11	0.15	1.00	-0.88	0.71	0.86	0.30	0.38	0.65	0.72	0.52
gyroY	0.05	0.23	0.09	-0.2	-0.33	-0.25	-0.59	-0.33	-0.36	-0.06	-0.10	-0.13	-0.30	-0.07	-0.10	-0.29	-0.27	-0.25	0.22	0.05	-0.58	1.00	0.05	-0.11	-0.3	-0.19	-0.15	-0.16	
gyroZ	-0.12	0.10	-0.06	-0.24	-0.20	-0.15	-0.36	-0.28	-0.06	-0.08	-0.10	-0.18	-0.20	-0.07	-0.08	-0.18	-0.27	0.06	-0.28	-0.28	1.00	0.46	0.07	0.21	0.46	0.53	0.53	0.29	
band.0	-0.07	-0.30	-0.23	-0.10	-0.1	-0.21	-0.10	0.03	0.46	0.22	0.10	0.10	0.31	-0.07	0.22	0.3	0.2	0.15	-0.27	0.17	0.66	-0.3	0.44	1.00	0.60	0.71	1.00	0.98	0.87
band.1	-0.03	-0.16	0.05	0.11	0.10	-0.0	0.25	0.16	0.43	0.25	0.12	0.05	0.2	0.05	0.11	0.31	0.23	0.17	-0.26	-0.02	0.30	-0.31	0.07	0.60	1.00	0.71	0.60	0.49	0.87
band.2	-0.13	-0.31	0.02	0.09	0.01	-0.18	0.07	0.15	0.42	0.2	0.10	0.11	0.3	-0.01	0.26	0.35	0.28	0.12	-0.12	0.0	0.38	-0.3	0.21	0.71	1.00	0.71	0.64	0.77	
band.3	-0.07	-0.30	0.23	-0.10	-0.1	-0.21	-0.10	0.03	0.45	0.22	0.10	0.11	0.30	-0.07	0.21	0.3	0.2	0.16	-0.27	0.16	0.65	-0.33	0.45	1.00	0.60	0.71	1.00	0.97	0.87
band.	-0.09	-0.31	0.29	-0.17	-0.20	-0.23	-0.1	-0.02	0.47	0.16	0.08	0.09	0.3	-0.08	0.22	0.37	0.2	0.13	-0.23	0.22	0.72	-0.35	0.53	0.98	0.49	0.64	0.97	1.00	0.78
band.	0.02	-0.27	0.08	0.10	0.07	-0.16	-0.11	0.17	0.43	0.29	0.15	0.11	0.2	0.01	0.17	0.30	0.27	0.18	-0.31	0.03	0.52	-0.36	0.29	0.87	0.77	0.87	0.78	1.00	

(b) 初心者の相関行列

三味線の状態	撥の位置/姿勢							筋電位								加速度/角速度				音声									
	sAng	dAng	bPosX	bPosY	bPosZ	bRotX	bRotZ	emg.1	emg.2	emg.3	emg.4	emg.5	emg.6	emg.7	emg.8	emgALL	acceX	acceY	acceZ	gyroX	gyroY	gyroZ	lendALL	band.0	band.1	band.2	band.3	band.	
sAng	1.00	0.11	-0.0	0.13	0.37	-0.40	0.01	0.21	-0.10	-0.10	-0.2	-0.08	-0.13	-0.19	-0.18	-0.26	-0.23	-0.18	0.25	0.36	0.22	-0.10	0.3	-0.21	0.0	-0.29	-0.21	-0.21	-0.17
dAng	0.11	1.00	-0.06	-0.24	-0.20	-0.15	0.05	0.27	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06
bPosX	-0.0	-0.19	1.00	0.25	-0.35	0.76	-0.37	-0.55	0.13	0.07	0.28	0.29	0.29	0.36	0.3	0.31	0.40	0.39	0.09	-0.25	-0.29	0.2	-0.45	0.59	0.22	0.60	0.59	0.51	0.41
bPosY	0.13	-0.64	0.26	1.00	0.29	0.06	-0.22	0.01	-0.02	-0.05	-0.12	0.01	0.0	0.12	0.07	0.00	-0.01	0.15	0.3	0.00	-0.50	0.45	-0.19	0.15	0.26	0.10	0.15	0.19	0.1
bPosZ	-0.37	-0.30	-0.78	0.29	1.00	-0.70	0.31	0.91	-0.1	-0.12	-0.33	-0.27	-0.20	-0.23	-0.22	-0.27	-0.37	-0.40	-0.01	0.59	-0.16	0.08	0.21	-0.43	0.00	-0.50	-0.43	-0.30	-0.2
bRotX	-0.40	0.10	0.76	0.06	-0.70	1.00	0.06	-0.65	0.11	0.10	0.28	0.18	0.13	0.21	0.2	0.21	0.30	0.38	0.01	-0.60	0.10	0.01	-0.28	0.44	0.09	0.53	0.42	0.32	0.23
bRotY	0.01	0.06	-0.37	-0.22	0.31	0.06	1.00	0.40	-0.06	-0.07	-0.07	-0.16	-0.06	-0.06	-0.03	-0.10	-0.12	-0.15	0.1	0.19	-0.28	0.17	-0.12	-0.13	-0.11	-0.12	-0.11	-0.08	
bRotZ	0.21	-0.22	-0.80	0.01	0.91	-0.65	0.40	1.00	-0.10	-0.07	-0.25	-0.25	-0.17	-0.20	-0.19	-0.19	-0.30	-0.49	-0.2	0.51	-0.12	0.0	0.12	-0.40	-0.0	-0.48	-0.41	-0.28	

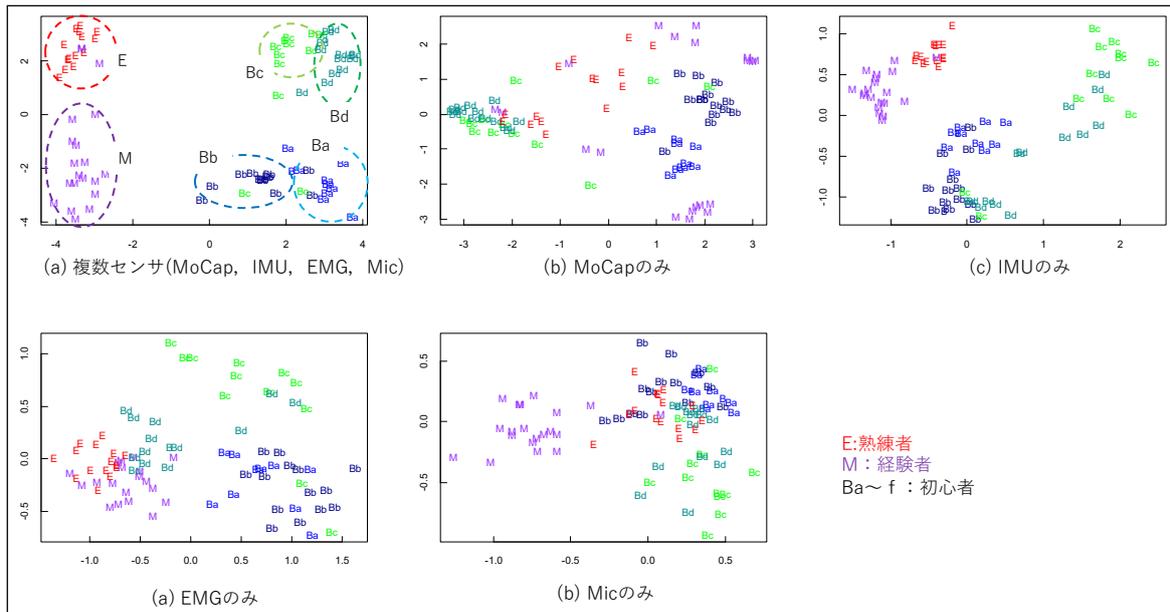


図7 演奏の類似性の可視化結果

(2) ニューラルネットワークによる分析

得られた相関行列について、演奏者毎の特徴を抽出できるか検討するために、得られた相関行列を入力とするCNNおよび2層の単純パーセプトロン (PT) を用いて演奏者毎に分類した。CNNのパラメータは、表1と同様で、学習回数は20000回とした。PTのパラメータは表3のとおり設定とした。演奏者毎のサンプル数が一定ではないので、学習データ24個、テストデータ12個(ただし学習データとは異なるデータ)を各クラスから抽出して分類する実験を100回繰り返し、分類率の平均を求めた。

表3 NNのパラメータ

学習回数	2000
学習率	0.01
活性化関数	softmax
勾配法	最急降下法
ミニバッチ	10

求めた平均および標準偏差を図8に示す。同図の結果から、CNNを用いた場合でもPTを用いた場合でも、複数のセンサを組み合わせた場合に、分類率が上がることがわかる。また、分類率が高い組み合わせであれば、CNNもPTも同程度の分類性であることがわかる。

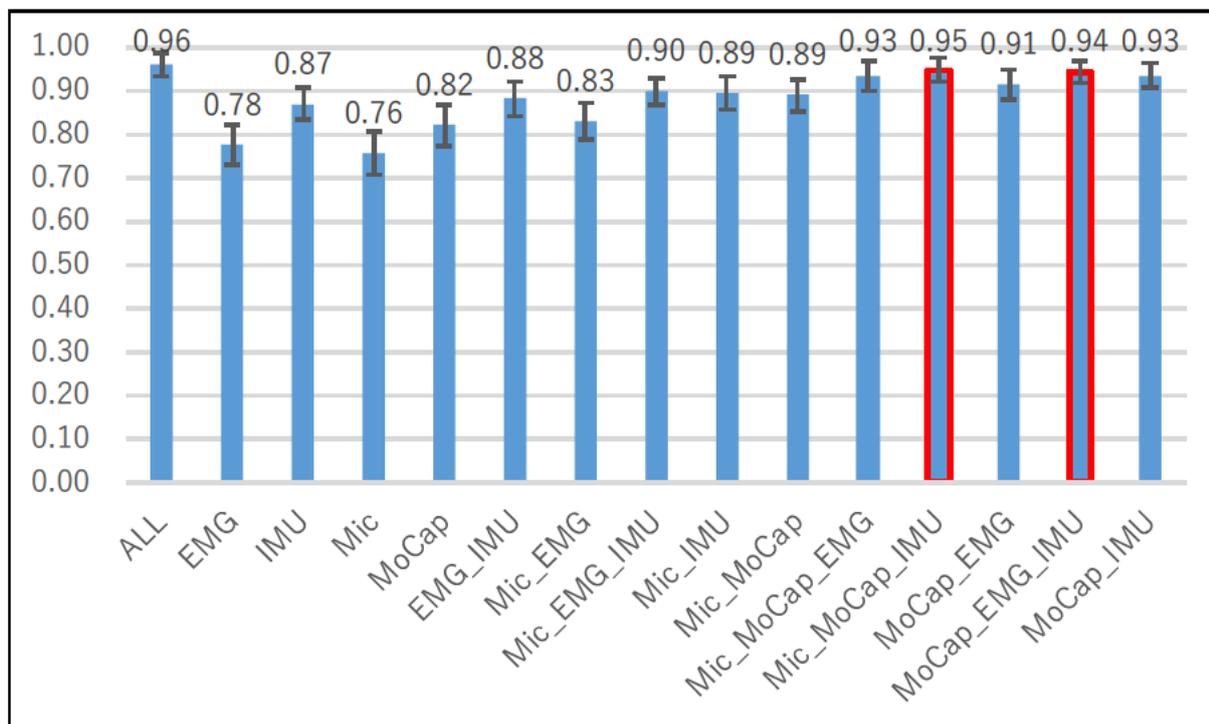


図 8 PT による分類結果

表 4 に、PT で抽出された各クラスの重みのうち、上位 5 つを示す。同表から、初心者は bPos や accel など動きに関するデータ間の相関の重みが大きくなっている。一方で熟練者は gyro と band の相関の重みが高い場合が多い。初心者が撥の位置や三味線の姿勢を意識しすぎて、意図せず同時に関連部位が動いてしまう（相関が高い）場合または、姿勢が安定しない（相関が低い）などの特徴が抽出された可能性がある。これに対して、熟練者は腕の gyro, accel など IMU で計測される腕の回転や加速度と band の相関が高い。発生する音とタイミングに合わせて腕の力加減を同期させるなど、熟練者のコツが表れている可能性が考えられる。

表 4 PT が学習した重み

初心者A			初心者B			熟練者		
bRotX	bRotZ	-0.16683	bPosY	bRotX	-0.20767	bPosZ	accelZ	-0.11931
bPosX	bRotX	0.127659	bPosZ	bRotX	-0.18154	gyroX	band.1.	0.118289
sAng	accelZ	0.122493	bPosX	bRotX	0.105617	bPosY	accelZ	-0.11561
bPosZ	bRotX	-0.12209	accelX	accelZ	-0.09801	gyroX	band.3.	0.109215
bPosX	bPosY	0.120512	bPosX	eABS.7.	0.096891	gyroX	band.ALL.	0.108227
初心者C			初心者D			経験者		
accelX	accelZ	0.143399	bPosZ	bRotX	0.15139	bPosY	band.1.	-0.11319
bPosZ	bRotX	0.137601	bPosY	bRotX	0.13505	accelZ	gyroY	-0.09767
accelY	gyroX	0.126641	dAng	bPosZ	0.129583	bPosZ	band.1.	-0.09184
accelZ	gyroZ	-0.1249	dAng	bPosY	0.127895	accelX	accelZ	-0.08865
gyroX	gyroZ	-0.12394	accelY	gyroZ	-0.12069	accelX	gyroY	0.087477

以上のことから、複数センサで計測した三味線演奏を NN によって分類し、分類結果を比較することによって、熟練者と初心者の演奏の傾向の違いを検討できることがわかった。また、本研究で得られた結果は、「肩の力を抜く」「よい姿勢を保つ」といった一般的に知られる知見と矛盾するものではなく、三味線演奏のコツについて、定量的な解析に関する知見が得られたといえる。

4 おわりに

身体技能の学習は、学習者が「コツ」として獲得し、繰り返し訓練することによって、自身の身体に最適化する。しかしながら、コツの言語化がされていない技能では、「見て学ぶ」といわれるように、学習者は観察と模倣によって自ら動作のコツをつかみ、動きを習得することとなる。本研究では、観察と模倣によって学習を進める技能として、日本の芸能において重要な和楽器である三味線に着目した。三味線の学習において、情報・電気通信技術を用いてコツを解析する手法について検討した。三味線の学習を進めるためには、動作の外見的な模倣にとどまらず、学習者が動きの特徴を理解すること、学習者が自身の動きを客観的にとらえて修正できることが重要になると考えられる。特に、身体動作と音の関連性を意識した学習が不可欠である。本研究は、1) 複数のセンサを用いて身体の動き、力の入れ具合、および音を計測するシステムの開発し、2) 計測した複数センサのデータを解析した。

実験では、熟練者、経験者、初心者の演奏を計測した。予備実験として、分割したデータをCNNによって楽曲、演奏経験のクラスに分類した。その後、PNおよびCNNを用いて個人毎に分類し、分類に寄与するデータについて考察した。本実験では熟練者は腕の慣性データと音に相関があることが特徴であり、初心者は撥と三味線の位置に相関が現れる傾向があることを示した。

本研究の成果は、熟練者のコツを学習者に示すことによって模倣を支援する、学習者の動きを解析してフィードバックし訓練を支援するなど、電気通信技術を用いた身体動作の学習支援への応用が期待できる。

【参考文献】

- [1] 柴田傑, 玉本英夫, 松本奈緒, 三浦武, 横山洋之: 学習者中心のインタラクティブ舞踊学習支援システムの開発, 電子情報通信学会論文誌 D, Vol.J97-D, No.5, pp.1014-1023, 2014.
- [2] 柴田傑, 玉本英夫, 海賀孝明, 横山洋之: 身体動作の3次元計測によるリアルタイム舞踊学習支援システム, 日本バーチャルリアリティ学会論文誌, 特集号 3次元インタラクション, Vol.17, No.4, pp.353-360, 2012.
- [3] Shibata T., Mitobe K., Miura T., Fujiwara K., Saito M., and Tamamoto H.: Development of an Uchi Self-learning System for Mutsumi-ryu-style Shamisen Using VR Environment, ICIRA 9th, pp.359-370, 2016.
- [4] 岩崎仁志, 柴田傑, プロジェクションマッピングを用いた三味線の構え方学習支援システムの提案, 情報処理学会インタラクション 2018, pp. 429-432, 2018.
- [5] 小坂谷壽一, 小玉成人, 川守田礼子: 伝統音楽(津軽三味線)保存用自動採譜装置に関する研究, 電子情報通信学会技術研究報告, 応用音響, Vol. 108, No. 491, pp. 37-42, 2009.
- [6] 濱中敬人, 坂本大介, 五十嵐健夫: Aibiki 譜面の適応的自動スクロールによる三味線演奏支援, 研究報告エンタテインメントコンピューティング 2014, No. 69, pp. 1-6, 2014.
- [7] 桂博章: 秋田県における民謡三味線の成立と流派間の奏法の違い, 秋田大学教育文化学部研究紀要, 文科学・社会科学部門, Vol. 70, pp. 1-7, 2015.
- [8] Perez-Carrillo A., Arcos J.-L. and Wanderley, M.: Estimation of guitar fingering and plucking controls based on multimodal analysis of motion, audio and musical score, International Symposium on Computer Music Multidisciplinary Research, Springer, pp.71-87, 2015.
- [9] Volpe G., Kolykhalova K., Volta E., Ghisio S., Waddell G., Albornò, P., Piana, S., Canepa C. and Ramirez-Melendez R.: A multimodal corpus for technology-enhanced learning of violin playing, Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter, ACM, p.25, 2017.
- [10] 津川信子: 三味線をはじめよう! 独習でも三味線で弾けるようになる, 成美堂出版, ISBN-13: 978-4415028323, 2004.
- [11] 岩田昌樹, 柴田傑: 畳みこみニューラルネットワークのためのモーションデータ画像化手法の提案, 第 21 回日本バーチャルリアリティ学会大会 大会論文集(21B05), pp. 1-4, 2016.

〈発 表 資 料〉

題 名	掲載誌・学会名等	発表年月
三味線演奏の特徴抽出に適した分割フレーム数の検討	情報処理北海道シンポジウム 2018 (22), pp. 1-2	2018 年 10 月
三味線演奏の複数センサを用いた計測システムの提案	じんもんこん 2017 論文集, pp. 69-74	2017 年 12 月
複数センサを用いた三味線演奏の特徴抽出手法の検討	第22回日本バーチャルリアリティ 学会大会 大会論文集 (1F303), pp. 1-4	2017 年 9 月

