

エレキギター演奏におけるミスの自動検出

下尾波輝[†] 矢谷浩司[†]

東京大学大学院工学系研究科電気系工学専攻[†]

1 はじめに

楽器を上手く演奏できるようになるには、多くの時間と労力を練習に費やす必要がある。広く親しまれている楽器の1つであるエレキギターにおいてもこのことは例外ではなく、エレキギター練習支援のためのシステムがこれまでに数多く開発されてきた。しかし、既存のエレキギター練習用システムは「正しいタイミングで正しい音程で弾けたか」のみしか演奏評価基準がない。そのため、例えば不適切な押弦によってノイズが混ざったような音が発せられたとしても、タイミングと音高さえ合っていれば正しく弾けていると見なされてしまう。

下尾と矢谷は、この問題点を解決するために、音色の綺麗さと演奏の滑らかさに着目し、単音のエレキギター演奏に自動で点数をつける研究を行った [1]。しかし演奏に対して点数を提示するのみでは、特に初心者は自分の演奏をどう改善すればよいのかわからない。上記の研究で5人の初心者にはエレキギターを演奏してもらった後に行ったインタビューでは、「左手の指を細かく動かすことが難しく、自分では上手く弾けているかの判別が難しい」という旨の発言が多く得られている。そこで本研究では、単音のフレーズのエレキギター演奏において、左手の押弦が適切に行われているかを判定してそれをユーザにフィードバックするシステムを提案する。

2 押弦におけるミスの分類

エレキギター経験が6年以上である著者1名を含む合計3名から意見を収集し、エレキギターで単音のフレーズを弾く際の押弦に関する主なミスを3つに分類した。

フレット上で弦を押さえている。

押弦する際にはフレット（金属製の棒）のすぐ近くを押さえるのが正しいが、フレット上で弦を押さえてしまうと不適切な音が出る。

十分に弦を押さえられていない。

フレットから離れた部分を押弦する、あるいは指の腹などで押弦するなどした場合、弦が十分に押さえられずに音がうまく鳴らない。

押弦している指を離すのが早すぎる。

押弦している指を離すと鳴っている音が止まるが、指を離すのが早すぎると音が滑らかに繋がらない。

このうち、3つめのミスの検出には下尾と矢谷がで考えた音の鳴り止め始めを検知する手法が使用できる [1]。したがって、本研究では残りの2つのミスに関する検知手法を議論する。特に、与えられた音を「正しく押弦できている音/フレット上を押さえている音/十分に弦を押さえられていない音」の3つに分類することで、これらのミスを検出することを目指す。

3 押弦におけるミスの自動検知手法

音から押弦に関係しそうな音響的特徴量を抽出し、その特徴量を用いて音を分類する。本節では、まず分析に用いるデータセットについて説明し、使用する音響的特徴量とそれを用いた分類性能について述べる。

3.1 データセット

音を「正しく押弦できている音/フレット上を押さえている音/十分に弦を押さえられていない音」の3つに分類する手法を考えるために、実際にこれらの音を鳴らしたデータセットが必要となる。そこでエレキギターの各弦（1~6弦）の1~12フレットで上記の3種類の音をそれぞれ3回ずつ鳴らしたデータセットを、第一著者が実際に演奏することで用意した。またエレキギターは3種類用いて、それぞれのエレキギターで上記の録音（24 bit/44100 Hz）を行った。そのため、録音した音は合計で1944個である。

3.2 使用する音響的特徴量

MFCC, Δ MFCC

音の振幅が最大になる時点周辺のメル周波数ケプストラム係数（MFCC）と、その時点から5フレーム（24 ms/frame, overlap=50%）における Δ MFCCを使用する。それぞれ1~12次の係数を用いる。

Tone Stability (TS_{np})

音の振幅が最大になる時点（peak）とそこから約46 ms後の時点（near_peak）におけるMFCC（1~12次）のcos類似度をTone Stability (TS_{np})とし、音色の安定度の指標として用いる [1]。

Harmonic-Percussive Ratio (HPR)

Harmonic-Percussive Source Separation (HPSS) [2] をエレキギターの音に適用すると、弾弦時の打楽器のような音とその後の調波音を分離することができる。分離した2つの成分について、それぞれの最大振幅の比の対数をHarmonic-Percussive Ratio (HPR) と定義する。

3.3 Support Vector Machine (SVM) による分類

上述した4種類の特徴量の組み合わせを変えながら、SVM (RBF カーネル, $C = 10.0$, $\gamma = 0.01$) による分類精度をcross validationによって比較する。cross validationのデータ分割方法には以下の2つを用いる。

Automatic Mistake Detection for Electric Guitar Performance Assessment

Namiki SHIMOO[†] and Koji YATANI[†]

[†] Interactive Intelligent Systems Laboratory,
Graduate School of Engineering, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
{shimoo, koji}@iis-lab.org

表1: 4 種類の特徴量の組み合わせを変化させたときの分類精度比較. leave-one-guitar-out cross validation においては, MFCC, Δ MFCC, HPR の 3 つを説明変数としたときの精度が最も高い.

Features				Mean ACC	
MFCC	Δ MFCC	TS _{np}	HPR	10-fold	LOGO
✓				.662	.562
	✓			.722	.709
		✓		.389	.384
			✓	.744	.709
✓	✓			.794	.689
✓		✓		.714	.620
✓			✓	.792	.723
	✓	✓		.722	.709
	✓		✓	.844	.820
		✓	✓	.730	.696
✓	✓	✓		.801	.702
✓	✓		✓	.879	.853
✓		✓	✓	.812	.739
	✓	✓	✓	.846	.828
✓	✓	✓	✓	.884	.844

表2: 3 種類の特徴量 (MFCC, Δ MFCC, HPR) を説明変数として LOGO を行ったときの分類結果.

		predicted			
		○	△	×	
actual	○	556	65	27	○... 正しい押弦
	△	70	533	45	△... フレット上を押弦
	×	14	64	570	×

10-fold cross validation (10-fold)

データをランダムに 10 分割にして, そのうち 9 個を学習データ, 残り 1 個をテストデータとする.

leave-one-guitar-out cross validation (LOGO)

2 つのエレキギターによる音を学習データ, 残り 1 つのエレキギターによる音をテストデータとする.

これら 2 種類の cross validation による性能評価の結果を表1に示す. 10-fold においては 4 種類の特徴量全てを用いたときの予測精度が最も高い ($ACC = .884$) が, LOGO においては TS_{np} を除いたときのほうが予測精度が高くなっている ($ACC = .853$). これは, エレキギターの個体差が TS_{np} に影響を与えているためであると推測される. そこで, 各エレキギターの「正しく押弦できている音」の TS_{np} について, 弦別 (1~6 弦) に分布を表示した (図1). その結果, エレキギターによっては低音側の弦の TS_{np} が低い傾向があることがわかった. 弦高などのコンディションにより, 低音側の弦に「ビビリ」が生じるなどしていることが原因であると考えられる.

4 練習支援システムの作成

前節で構築した音の分類モデルと以前の研究 [1] で構築した演奏の自動採点モデルを用いて, エレキギターの練習を支援するシステムのプロトタイプを作成した (図2). 4 つの 8 分音符 (BPM100) から成るフレーズを弾くと, 0~100 の点数が表示され, さらに各音について

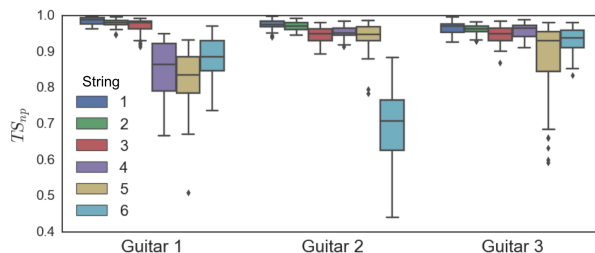


図1: 各エレキギターの正しく押弦できた音の弦ごとの TS_{np} 分布.



図2: 作成したシステム. 指定したフレーズ (4 音) を弾くと, その演奏の点数と各音の押弦の適切さを判定する. この図は, 指定したフレーズを 5 回繰り返して演奏したときの結果を表している.

押弦が適切に行われていたかの判定が行われる. このシステムを用いることにより, 既存システムでは考慮されていなかった音色の綺麗さなどにも注意しながら練習できることが期待される.

5 おわりに

本研究では, エレキギター初心者が特に難しいと感じる左手の指による押弦に焦点を当て, 音を「正しく押弦できている音/フレット上を押さえている音/十分に弦を押さえられていない音」の 3 つに分類する手法を考案した. そしてその手法と以前の研究の結果を用いて, エレキギター演奏に自動で点数をつけ, 押弦が適切に行われていたかを判定するシステムを作成した. ユーザスタディによって作成したシステムの有用性の検証を行うことが今後の課題である. また, 今回は 3 つのエレキギターおよび単一の演奏者のみの演奏音で分析を行った. そのため, 多様な演奏者とエレキギターによるデータを増やし, 今回の分析結果の汎用性を確認する必要がある.

参考文献

- [1] 下尾波輝, 矢谷浩司: エレキギター演奏自動評価のための音響的特徴量の調査, 研究報告音楽情報科学 (MUS), Vol. 2017-MUS-117, No. 3, pp. 1-6 (2017).
- [2] Fitzgerald, D.: Harmonic/Percussive Separation using Median Filtering. *Proc. DAFX-10*, pp.217-220 (2010).