

# ギター演奏者の習熟度に合わせた音響信号 からのタブ譜自動生成

矢澤 一樹 糸山 克寿 奥乃 博

山内 拓海

情報システム工学科

2023年7月7日

## 背景

タブ譜には押さえるべき弦とフレットの位置が数字によって表されており、音楽的知識の乏しい人でも直感的に演奏を行うことができる。しかしこのような重要性に関わらず、タブ譜の数は一般的な五線譜の数に比べて少ない。

## 本研究の目的

演奏支援への応用を考えた場合、初心者から熟練者まで多様なギター演奏者のニーズに対応するためには、演奏者の習熟度に合わせたタブ譜を生成すること

## 提案手法の概要

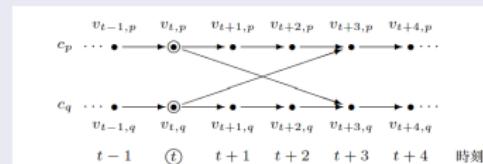
既存の多重基本周波数推定法を用いて音高推定を行い、その後その結果を基に運指推定を行う。さらに、推定された最適運指を用いて後処理を行うことで、元の音高推定結果に含まれるギターで演奏不可能な音高の組み合わせを排除する

## 潜在的調波配分法 (LHA)

LHA は機械学習を用いた多重基本周波数推定法の一つである。楽器音の調波構造をネスト型混合ガウス分布によって近似し、分布の各パラメータをベイズ推定することで音高の推定を行う。  
以降では、LHA の推定結果として得られた時刻  $t$  における音高  $k$  の出現度を、 $N_{tk}$  と表す。

## 運指推定と音高推定結果への後処理

ギター演奏における運指を押弦フォームの時間的推移とみなし、これを図 1 に示す重み付き有向グラフでモデル化する。  
最適フォーム系列の推定後、各フォーム  $c_{pt}$  で演奏不可能な全ての音高を、元の音高推定結果から排除する。



# 重み付き有向グラフの詳細

4/11

## 重み付き有向グラフの詳細

本手法では各辺の重みを音響再現度と運指容易度の重み付け和によって定義する。

これにより各演奏者の習熟度に応じた運指を推定可能にした。  
ギター演奏に関する 3 つの制約に基づいて決定した。

## 制約 1: 押弦可能フォーム制約

楽曲全体の運指が演奏可能であるためには、各時刻で用いられるギターフォームが押弦可能であることが不可欠である。

ギターのコード表を基に、ギター演奏で用いられるフォームのテンプレートを全て列挙する。(総数 P=1401)

## 制約 2: フォーム変化時刻制約

入力音に対し発音時刻検出を行い、検出された発音時刻でのみフォームを変更できるという制約を加える。

発音時刻は、以下の値 (Ntkflux: NF) を用いて検出する。

$$N_{Ft} = \sum_k \max(0, N_{tk} - N_{(t-1)k})$$

## 制約 3: 同一フォーム継続制約

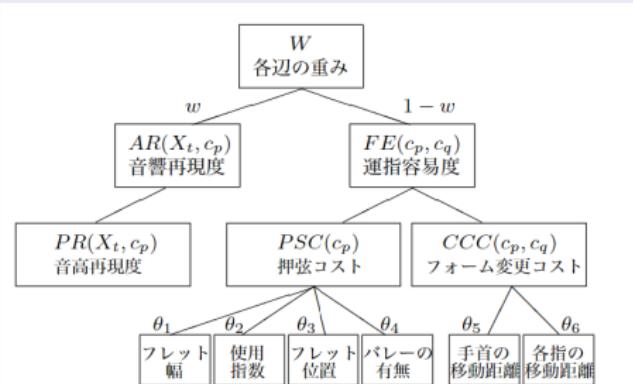
フォーム変化後最低 D 時間の間は同一のフォームを用い続けなければならないという制約を加える。

最低継続時間 D は押弦フォームの種類やフォーム変化時刻によらず、楽曲全体で同一の値であると仮定する。

## 各辺の重み

グラフの各辺  $e_{tupq}$  の重みは、音響再現度 (Acoustic Reproducibility: AR) と運指容易度 (Fingering Easiness: FE) に基づいて定義する。パラメータ  $w$  を変更することで演奏者の習熟度を運指推定に反映できるようにする。

$$W_{tupq} = u \sum_{t'=t+1} (w^* \cdot AR(X_{t'}, c_{pt'}) + (1 - w) \cdot FE(c_{pt'}, c_{pt'+1}))$$



## 音響再現度 (PR)

音響再現度は、ある運指が与えられたときに、その運指によって実際の演奏音がどの程度再現されるかを表す。

音響再現度として音高再現度のみを考慮する。押弦フォーム  $c_p$  による音高  $X_t$  の再現度 (Pitch Reproducibility: PR) を、以下の式で定義する

$$PR_{(X_t, c_p)} = \sum_{k \in K_p} N_{tk}$$

## 運指容易度 (FE)

運指容易度は、各フォームがどれだけ押弦しやすいかという押弦コスト (PSC) と、各フォーム間の変更がどれだけ行いやすいかというフォーム変更コスト (CCC) によって決まる。

$$FE(c_p, c_q) = \frac{1}{1 + (PSC(c_q) + CCC(c_p, c_q))}$$

## 押弦コスト (PSC)

- a1 : フレット幅
- a2 : 使用指数
- a3 : フレット位置
- a4 : バレーの有無

$$P_{SC}(c_p) = \frac{\sum_{i=1}^4 \theta_i a_i(c_p)}{\max_p \left( \sum_{i=1}^4 \theta_i a_i(c_p) \right)}$$

## フォーム変更コスト (CCC)

本手法では、手首の移動を考慮したフォーム変更コストを、以下の手順で計算する。

- (1) フォームの押弦に用いている全ての指を同じ距離だけ水平移動ときの水平方向の移動距離を、手首の移動距離 a5 とする。
- (2) 2 つのフォーム間での各指のマンハッタン距離を計算し、その総和を a6 とする。指の追加および離弦に対しては指の種類によらずマンハッタン距離を 1 とする。

$$CCC(c_p, c_q) = \frac{\sum_{i=5}^6 \theta_i a_i(c_p, c_q)}{\max_{p,q} \left( \sum_{i=5}^6 \theta_i a_i(c_p, c_q) \right)}$$

## 実験条件

実験データは RWC 音楽データベースのジャズ楽曲 9 曲およびポピュラー楽曲 52 曲 (冒頭 60 秒)

同一フォーム最低継続時間  $D=200\text{ms}$

個人の運指傾向を反映するためのパラメータ  $\theta$  は,  $\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6 = 4, 4, 1, 4, 2, 1$

## 評価基準

音高推定精度の評価基準には, 時間フレーム単位での適合率, 再現率,  $F$  値を用いた.

運指難易度の評価には, 各パート中に出現する押弦フォームの総数および, 本稿で定めた運指コストのパート全体に対する総数を用いた.

## 実験結果

・  $w$  の値を大きくして音響再現度を重視したときには、 $F$  値が向上  
 → LHA の推定結果に含まれる不適切な音高が、提案手法で用いた制約によってある程度排除できていることを示す。

表 2 および図 3 より、 $w$  の値を小さくするほど出現フォーム数や運指コストが少なくなり、運指の容易な初心者用のタブ譜が出力されることが確認された。

表 1 提案法および LHA の音高推定結果。値は全パートに対する平均値を表す。

評価基準	LHA						提案法					
	$w = 1.0$	0.90	0.80	0.70	0.60	0.50	$w = 1.0$	0.94	0.743	0.742	0.730	0.712
F 約	0.702	0.744	0.743	0.742	0.730	0.712	0.679					
適合率	0.707	0.754	0.761	0.766	0.777	0.786	0.780					
再現率	0.704	0.743	0.734	0.727	0.699	0.659	0.655					

表 2 提案法の運指推定結果。値はパートに対する平均値を表す。

評価基準	$w = 1.0$	0.90	0.80	0.70	0.60	0.50
出現フォーム数	29.6	29.4	27.4	25.1	22.6	18.8
運指コスト	1014.1	713.9	603.1	489.0	389.6	312.0



(a)  $w = 1.00$



(c)  $w = 0.50$

(a) 表 1, 2 図 3

## 考察

本実験ではギターに特殊な装置を装着しないで従来法よりも多い1000 以上の押弦フォームにさらに本手法は、個人の演奏レベルに応じた運指を推定可能となった。

## 今後の課題

- θ 変更時に推定される運指の妥当性や各項目の重要度などを今後調べる必要がある。
- 音高再現度と運指容易度の両面からタブ譜を評価したが、タブ譜そのものの推定精度について評価する。