

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^{1,a)} 糸山 克寿^{1,b)} 奥乃 博^{1,c)}

概要：本稿では、ギター演奏者の演奏支援をするために実際のギター演奏音から演奏者の習熟度に応じたタブ譜を自動生成する手法について述べる。具体的には、初級者向けには音符の欠落などを許容してでも演奏が容易なタブ譜を、上級者向けには音高を正確に再現するタブ譜を、それぞれ生成する。推定される運指の難易度は、音響再現度と運指容易度の相対的な重みをユーザー側で調整することによって変更可能である。本手法によって得られたタブ譜について音響再現度と運指容易度の両面から評価を行った結果、パラメータを変更することによって音高推定の適合率を保ったまま運指を簡略化できることが確認された。

1. はじめに

タブ譜は、ギターやベースなどの弦楽器用の楽譜形式の一つであり、特に多くのギター演奏者にとって馴染み深いものである。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。また、ギターは異なる弦で同一の音高を演奏することができる楽器であり、そのため一つの音高列に対して複数の運指の可能性が考えられる。タブ譜を用いることで、演奏者はこのような運指の多様性に悩まされることなく、簡単に演奏を行うことができる。このようにタブ譜はギター演奏者、とりわけギター初心者にとって、重要な練習材料となっている。

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。最近では、CGM(Consumer Generated Media)の普及によって、自分の演奏したい楽曲のタブ譜を入手するのがますます難しくなっている。さらに、Web上にあるタブ譜には間違いも多く含まれており、正確なタブ譜を簡単に検索することは難しい。Webから高精度なタブ譜を検索する手法 [1] なども提案されているが、彼らの手法は同一曲に対する複数のタブ譜同士の一貫性を特徴量として用いるため、タブ譜の数が少ないマイナー曲などに対して利用することはできない。以上のような理由から、ギター演奏者の演奏支援のためには、音響信号からのタブ譜自動生成システムの実現が望ましい。

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が重要である。一般に、ギター演奏者は耳コピなどによって運指を決定する際、音響再現度と運指容易度の両方を考慮すると言われている。ここで、音響再現度とはある運指によって演奏される音高や音色が実際の演奏音にどの程度近いかを、運指容易度とは演奏者にとってその運指がどの程度演奏しやすいかを、それぞれ表す。両者はトレードオフの関係になっており、どちらを重視するかは演奏者の演奏習熟度によって異なる。すなわち、ギターの熟練者は多少演奏が難しくてもより実際の演奏に即した音が出せるような運指を行い、逆に初心者は多少音符の欠落などを許容してでも演奏が容易な運指を選択する傾向がある。さらに、演奏レベルが同程度の演奏者間でも、細かな演奏傾向や苦手な演奏法などは個人によって異なる。例えば、バレーコードと呼ばれる押弦法が苦手な場合や、あるいは手が小さいために指を大きく広げるフォームが演奏し難い場合などが考えられる。このような知見を考慮すると、演奏者の習熟度や演奏傾向に応じて異なるタブ譜を生成することが望ましい。

本稿では、音響信号からギター演奏者の習熟度に応じたタブ譜を自動生成するシステムを提案する。我々は、既存の多重基本周波数推定法 LHA を用いて各時刻における各音高の出現度を推定し、その後得られた音高出現度と各押弦フォームの演奏コストに基づいて最適な運指を推定する。提案法では、運指推定を重み付き有向グラフ上での最長経路探索問題としてモデル化し、これを動的計画法を用いて解くことで、ギター演奏時の身体的制約を考慮した運指推定を行うことができる。またこのとき推定される運指

¹ 京都大学
Kyoto University, Kyoto-shi, Kyoto 606-8501, Japan
a) kyazawa@kuis.kyoto-u.ac.jp
b) itoyama@kuis.kyoto-u.ac.jp
c) okuno@kuis.kyoto-u.ac.jp

は、音響再現度と運指容易度の相対的な重みを変更することで、演奏者の習熟度に応じて調整可能である。

本稿の構成は以下のとおりである。第二章で音高推定と運指推定の概要を述べ、続く第三章で運指推定で用いる重み付き有向グラフの詳細について述べる。第四章では、本システムを評価するために行った実験とその結果について述べる。最後に、関連研究との比較および今後の課題について述べ、本研究のまとめを行う。

2. 提案法の概要

提案法では、初めに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。本章では、これら音高推定と運指推定の概要について述べる。なお本手法の流れそのものは、我々の先行研究 [2] とほぼ同じである。

2.1 LHA を用いた音高推定

初めに、潜在的調波配分法 (Latent Harmonic Allocation: LHA) [3] を用いて音高推定を行う。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。

LHA では、音響信号の周波数スペクトルを入力として与えることで、出力として各時刻における各音高の出現度が得られる。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} と表すこととする。

2.2 運指推定と音高推定結果への後処理

次に、LHA による音高推定結果を基に最適な運指の推定を行う。本手法では、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。グラフの頂点や辺の有無はギター演奏に関する 3 つの制約に基づいて決定し、各辺の重みは音響再現度と運指容易度によって定義する。これらの制約や重みの詳細は、次章で述べることとする。本グラフの最長経路を動的計画法を用いて探索することで、最適押弦フォーム系列 $C^* = \{c_{p_1}^*, \dots, c_{p_T}^*\}$ が得られる。

最適フォーム系列の推定後、各フォーム $c_{p_t}^*$ 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。時刻 t での音高 k の修正後の音高出現度を \tilde{N}_{tk} とすると、 \tilde{N}_{tk} は以下のように定義される。

$$\tilde{N}_{tk} = \begin{cases} N_{tk} & (k \in K_{p_t}) \\ 0 & (\text{otherwise}) \end{cases}$$

ここで $K_{p_t} = \{k_{p_t1}, \dots, k_{p_t6}\}$ は、最適押弦フォーム $c_{p_t}^*$ の 6 つの弦で演奏可能な音高の組み合わせを表す。

押弦可能フォーム

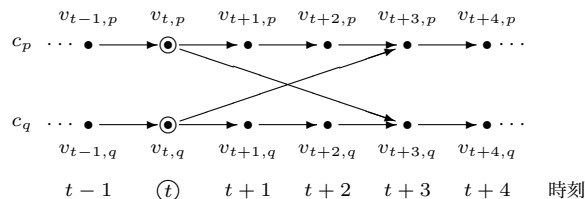


図 1 最適運指推定は、重み付き有向グラフ上での最長経路探索問題としてモデル化できる。本図は押弦可能フォーム数 2、 $D = 3$ で、時刻 t のみが発音時刻の例である。

2.3 タブ譜生成

最終的な音高の有無は、 \tilde{N}_{tk} を閾値判定することで決定する。すなわち、閾値パラメータを α とし、 $\tilde{N}_{tk} > \alpha \max \tilde{N}_{tk}$ を満たす全ての音高 k を時刻 t で鳴っている音高とみなす。

提案法では、こうして推定された最適押弦フォーム系列 C^* および音高出現度 \tilde{N}_{tk} の閾値判定結果を用いることで、各時刻で演奏された弦およびフレットを求めることができる。さらに、これらの情報と既存の拍検出手法 [4] や発音時刻検出手法 [5] による推定結果などを組み合わせることで、実際のタブ譜を生成することができる。

3. 重み付き有向グラフの詳細

本章では、運指推定に用いる重み付き有向グラフの定式化を行い、頂点や辺の有無、各辺の重みの決定方法について述べる。先行研究 [2] と異なり、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。これにより我々は、各演奏者の習熟度に応じた運指を推定可能にした。また頂点や辺の有無を決定するギター演奏に関する 3 つの制約についても、本稿で若干の修正を行った。

図 1 において、グラフの縦軸は押弦フォームの種類、横軸は時刻をそれぞれ表し、各有向辺はフォームの推移を表す。以降では、 v_{tp} を時刻 t でのフォーム c_p の使用を表す頂点とし、 e_{tupq} を頂点 v_{tp} から v_{uq} への有向辺とする。本グラフの最長経路が通る各頂点 $\{v_{1p_1}^*, \dots, v_{T p_T}^*\}$ が、最適フォーム系列 $C^* = \{c_{p_1}^*, \dots, c_{p_T}^*\}$ に対応する。

3.1 3 つの制約

グラフの頂点および辺の有無は、ギター演奏に関する 3 つの制約に基づいて決定する。以降で、それぞれの制約の詳細について述べる。

3.1.1 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。そこで提案法では、押弦可能なギターフォームをあらかじめ全て列挙しておき、各時刻での最適フォームをこれらの押弦可能フォームの中から一つ選ぶこととする。

押弦可能フォームの列挙は、次の手順で行う。まず、ギターのコード表 [6] 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。ここでフォームのテンプレートとは、コード表に登場する各フォームから指の相対的な位置関係の情報のみを抽出したものである。その後、これらのテンプレートをギター上の任意の位置に配置することで得られる全ての押弦フォームを、押弦可能フォームとして列挙する。列挙された押弦可能フォームの総数は、 $P = 1401$ となった。図1の各頂点 v_{t1}, \dots, v_{tP} にこれらの押弦可能フォームをそれぞれ対応させることで、各時刻における押弦可能性を保障することができる。

3.1.2 フォーム変化時刻制約

演奏中の押弦フォームの変更は通常、その時点のフォームで演奏不可能な音高を演奏するために行われる。そこで我々は、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、spectral flux [5] を応用した以下の値 (N_{tk} flux: NF) を用いて検出する。

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

ここで、 N_{tk} は LHA の出力として得られた時刻 t での音高 k の出現度を表わす。 NF_t が特定の閾値 $\beta \max_t NF_t$ よりも大きくなる全ての時刻 t を発音時刻とみなす。

フォーム p からフォーム q へのフォーム変化は、図1のグラフ上では辺 $e_{\hat{t}(t+D)pq}$ で表される。ここで \hat{t} は、上記の方法により検出された発音時刻のうちの一つである。また、 D は以降で述べる最低フォーム継続時間を表す。

3.1.3 同一フォーム継続制約

人間が指を動かせる速度には限界があるため、フォーム変化が頻繁に起こりすぎるようなタブ譜は不適切である。そこで我々は、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。

図1のグラフ上の辺 $e_{\hat{t}(t+D)pq}$ は、時刻 \hat{t} で使用フォームがフォーム c_p からフォーム c_q に変化し、さらにその後 D 時間の間フォーム c_q を使用し続けることを表す。また本稿では、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。

3.2 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 Easiess: FE) に基づいて定義する。すでに述べたように両者は一般にトレードオフの関係にあり、運指決定の際にどちらを優先するかは演奏者の習熟度によって異なる。そこで提案法では、グラフの重みを以下のように音響再現度と運指容易度の重み付き和として定義し、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。

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

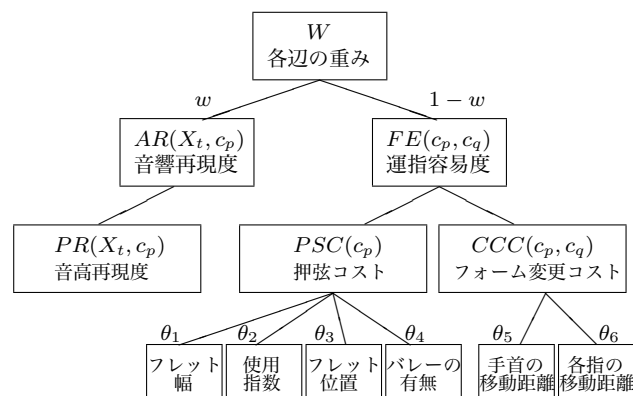


図2 グラフの重みの階層図。グラフの重みは、音響再現度と運指容易度の重み付き和として定義され、それぞれの項目はさらに下位の項目の組合せによって定義される。

パラメータ w の値が小さくなるほど、音響再現度に対する運指容易度の重要度が高くなり、より初心者向けの運指が推定されることが期待される。逆に w の値が大きくなるほど、運指容易度に対する音響再現度の重要度が高くなり、 $w = 1.0$ のときには運指容易度を全く考慮せず音響再現度を最大限に優先するような運指が選ばれることとなる。

音響再現度および運指容易度は、図2に示す階層構造で定義する。以降で、それぞれの項目の詳細を説明する。

3.2.1 音響再現度

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。ここで音響を特徴づける要因としては、主に音高や音色が考えられる。音高は、周波数スペクトルにおける基本周波数に対応しており、運指が一つ決まれば各時刻で演奏可能な音高の組み合わせは一通りに決まる。また音色は、音響信号上では調波構造の倍音比に対応しており、運指においては使用する弦の種類に対応している。すなわち、ギターでは異なる弦で同一の音高の音を演奏することができるが、それぞれの弦が発する音は音色が微妙に異なっており、これを考慮することでより実際の演奏の音色に近い音を出すことが可能である。しかしバイオリン演奏 [7] などに比べ、ギター演奏ではこのような弦の種類による音色の違いを考慮することは少なく、運指決定の際には音色を厳密に再現することよりも運指の容易さを重視する傾向が強い。そこで本稿では、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。

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

すなわち、あるフォーム c_p による音高 X_t の再現度は、そのフォームで演奏可能な6つの音高 $K_p = \{k_{p1}, \dots, k_{p6}\}$ の時刻 t での出現度の総和として定義される。

3.3 運指容易度

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容易度と、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容易度によって決まると考えられる [8,9]。そこで本稿では、各フォームに対する押弦コスト (Press down Strings Cost: PSC) とフォーム変更コスト (Configuration Change Cost: CCC) をあらかじめ計算し、これらを基に運指難易度を以下のように定める。

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

押弦コスト PSC は、運指推定に関する先行研究 [8-10] を参考に、以下の 4 つの項目によって定義する。

- a_1 : フレット幅
- a_2 : 使用指数
- a_3 : フレット位置
- a_4 : バレーの有無

フレット幅 a_1 は指を広げる幅であり、使用指数 a_2 は押弦に用いる指の総数である。またフレット位置 a_3 は、人差し指が押弦している弦のフレット番号とする。ギター指板はフレット番号が高くなるほどフレットの幅が狭くなるため、この値が大きくなるほど押弦が難しくなると考えられる。バレーの有無 a_4 については、一般的にバレーコードのほうがオープンコードに比べ押弦が難しい傾向があるため、押弦フォームがバレーコードの場合は $a_4 = 1$ 、オープンコードの場合は $a_4 = 0$ とする。押弦コスト PSC は、これら 4 つの項目の重み付け和によって定義される。

$$PSC(c_p) = \frac{\sum_{i=1}^4 \theta_i a_i(c_p)}{\max_p (\sum_{i=1}^4 \theta_i a_i(c_p))}$$

ここで θ は、4 つの項目の相対的な重みを決定するパラメータであり、この値を変更することで個々の演奏者の演奏傾向を運指推定に反映することができる。例えば、手が小さく指を広げるのが困難な演奏者に対しては θ_1 の値を相対的に大きくすることで、またバレーコードを演奏するのが苦手な演奏者に対しては θ_4 の値を大きくすることで、各演奏者にとってより負荷の小さい運指が推定されると考えられる。また、押弦コストとフォーム変更コストの影響度を等しくするため、全フォームに対する押弦コストの最大値が 1 となるように正規化しておく。

次に、二つのフォーム間でのフォーム変更コスト CCC を定義する。先行研究 [11,12] では、フォーム変更に必要なコストは各指について両フレット間のマンハッタン距離を計算し、それらを総和することで計算される。しかし実際のギター演奏を考えた場合、各指は独立に動くわけではないため、上記のようなコストの定め方では不十分であると考えられる。そこで本手法では、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) 両フォームのフレット位置が等しくなるように、一方

の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する。このときの水平方向の移動距離を、手首の移動距離 a_5 とする。

(2) 手首の移動後、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 a_6 とする。ただし、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。

(3) フォーム変更コストを、手首の移動距離 a_5 と、各指の移動距離の総和 a_6 の重み付き和として定義する。

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

ここでも θ の値を変更することで、各演奏者の演奏傾向を反映することができる。

4. 評価実験

本手法の性能を評価するため、パラメータ w を変更した時の、音高推定精度と運指容易度を調べる実験を行った。

4.1 実験条件

実験データには、RWC 音楽データベース [13] のジャズ楽曲 9 曲およびポピュラー楽曲 52 曲から抽出した、計 79 のギターパートを用いた。計算時間削減のため、各パートの冒頭 60 秒のみを実験に用いた。音響信号は MIDI シンセサイザー (YAMAHA MOTIF-XS) を用いて録音し、20ms の時間分解能でウェーブレット変換した。音高列の正解データは対応する MIDI データから作成した。

音高推定時の閾値パラメータ α は、各パートおよび各条件に対して F 値が最大となるように最適化した。発音時刻検出の閾値パラメータ β は、我々の先行研究 [2] を参考に $\beta = 0.30$ と定めた。また同一フォーム最低継続時間 D については、一般的なギター演奏者が 1 秒間に 5 回以上フォーム変更を行うのは難しいと考え、 $D = 200$ (ms) とした。個人の運指傾向を反映するためのパラメータ θ は、今回は実験的に $\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6 = \{4, 4, 1, 4, 2, 1\}$ とした。

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。比較のため、同一の実験データを用いて従来法である LHA の音高推定精度の評価も行った。ま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。

4.2 実験結果

提案法および LHA の音高推定結果を、表 1 に示す。実験結果より、 w の値を大きくして音響再現度を重視したときには、従来の LHA に比べて F 値が向上することが確認された。このことは、LHA の推定結果に含まれる不適切な (実際には演奏されていない) 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。ま

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
F 値	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	0.789
再現率	0.704	0.743	0.734	0.727	0.699	0.659	0.605

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

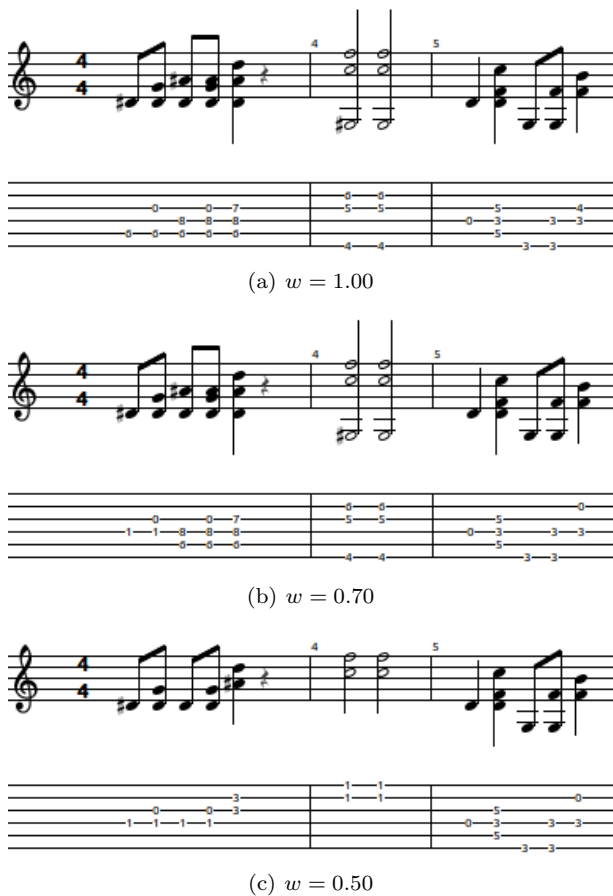


図 3 出力タブ譜の実例 (楽曲: RM-J007)

表 1 より, w の値が小さくするにつれて, F 値の値が下がっていくことが分かる. これは, 運指容易度を重視して音符の欠落などを許容した結果, 再現率が下がってしまうからであると考えられる. 一方適合率は, w を 0.50 まで下げても, 従来の LHA よりも高い値を保持できていることが分かった. このことは, 運指容易度を重視した初心者用のタブ譜において, 音高の推定誤りが少ないことを示す.

表 2 に, 提案法によって推定された運指の出現フォーム数と運指コストを示す. また, 図 3 に出力されたタブ譜の実例を示す. 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.

5. 考察

5.1 関連研究との比較

音響信号からのタブ譜自動生成に関する研究として, これまでに HMM を用いた手法 [11] や Inharmonicity に着目した手法 [14], NMF を用いた手法 [15], 視覚情報を併用した手法 [16] などが提案されている. しかしこれらの手法には, 推定可能なギターフォーム数や演奏法に限られている, ギターに特殊な装置を装着しなければならない, 音響信号の他にビデオデータが必要である, といった問題がある. それに対し我々の手法は, 本手法は入力として音響信号のみしか用いず, コード演奏やアルペジオ, ギターソロといったどのようなギター演奏法に対しても用いることができ, 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であり, その点が上記の手法とは大きく異なる.

ギターの運指推定に関して, 演奏習熟度や個人の演奏傾向を考慮した先行研究も存在する. 江村ら [10] は, ギター演奏時の各動作の負荷値を実演奏時のミス量の量から最小 2 乗法によって計算し, 演奏者にとって最も演奏しやすい押弦フォーム列を推定する手法を提案している. 藤井ら [17] は, 物理シミュレータ上で手の移動距離やトルクを算出することで最適な運指を推定する手法を提案しており, 本手法は手のモデルの大きさを変えることで各個人の指の長さに応じた運指を推定できる. これらの手法は個人の演奏傾向については考慮しているものの, 初心者・熟練者といった一般的な演奏習熟度についてはほとんど考慮していない. また入力として正確な音高列やコード列が与えられることを仮定しており, 本研究のような音響信号からの採譜に直接用いることはできない. しかし上記の手法は, 個人の演奏傾向を運指推定に自動で反映させる仕組みを作った点や, 運指の難易度に影響を与える因子を発見したという点で有意義であり, 今後我々の手法にも応用させていきたい.

5.2 今後の課題

5.2.1 追加実験

本稿では, 演奏者の習熟度を反映するパラメータ w を変更した時のシステムの性能を評価したが, 個人の演奏傾向を反映するパラメータ θ については評価を行わなかった. そこで, 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. また先行研究 [10] のように, 個人の実演奏データを用いて演奏者ごとの θ や w の最適値を求めることで, ユーザー側の負担を減らすことができると思われる. 演奏難易度が同程度の複数のタブ譜を用いて w の最適値をあらかじめ推定することなども考えられ, 今後はそれらの定式化を進めていきたい.

また本稿では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を

行わなかった。これは、RWC 楽曲には正解データとなるタブ譜が存在しないためである。しかし、もし同一曲に対する難易度の異なる複数のタブ譜が得られれば、 w 変更時に推定される運指の妥当性を定量的に評価できる。そこで今後、そのようなデータの収集や実験の実施を行いたい。

5.2.2 推定精度の向上

運指推定で用いた有向グラフの重みの決定法やギター演奏に関する制約の改善を行うことで、本手法には性能の向上の余地があると期待される。例えば音高再現度に関して、先行研究 [18, 19] では楽器音の調波構造を事前情報として与えることで多重基本周波数推定の精度が向上することが確認されており、ギターの調波構造の情報を本手法のモデルに組み込むことで、音高推定精度が向上することが期待される。また、楽曲中のメロディラインやコードの種類などをあらかじめ強調・推定することで、楽曲の響きのある程度保ったまま運指を簡略化することなども考えられる。

運指推定の面では、既存のテンポ推定法などを用いてあらかじめテンポを推定し、その情報をフォーム変化時刻制約や同一フォーム継続制約に利用することが考えられる。また押弦コストやフォーム変化コストに加えて、各フォームの統計的な出現度やツーファイブなどの音楽的なコードの推移傾向をモデルに組み込むことで、音楽的により妥当性の高いフォーム推移を推定することが可能になると思われる。先行研究 [10, 17] では、手のトルクや押弦に用いる指の違いが運指の難易度に影響を与えるということも知られており、それらの項目の検討も今後進めていきたい。

5.2.3 演奏支援以外への応用

本研究では、タブ譜自動生成をギター演奏者の演奏支援に用いることを目指したが、本システムは演奏支援以外にも応用が可能である。例えば、運指推定と編曲は共通の枠組みで扱えることが知られており [20]、本システムを応用することでピアノの演奏音をギターで演奏可能な形にアレンジすることなどができると考えられる。また本手法は押弦可能フォームの列挙法や運指コストの定義を変更することで、ギター以外の楽器の採譜にも利用することができる。

さらに提案法では、同一曲に対する音響信号とタブ譜のデータが得られた際、それらのデータを用いてパラメータ w の最適値を逆推定することでそのタブ譜の難易度を推定することが可能であり、演奏難易度に基づく音楽検索などにも応用できると思われる。本研究の今後の方向性の一つとして、検討していきたい。

6. おわりに

我々は、音響再現度と運指容易度の相対重みを変更することで、ギター演奏者の習熟度に応じたタブ譜を自動生成する手法を提案した。今後はパラメータの自動調整法の考案や、運指推定に用いるグラフの重みの各項目の検討、演奏支援以外の分野への応用などを行っていきたい。なお、

本研究の一部は科研費 No.24220006 (S) の支援を受けた。

参考文献

- [1] Macrae, R. and Dixon, S.: Guitar tab mining, analysis and ranking, *Proc. ISMIR*, pp. 453–458 (2011).
- [2] Yazawa, K. et al.: Audio-Based Guitar Tablature Transcription using Multipitch Analysis and Playability Constraints, *Proc. ICASSP*, pp. 196–200 (2013).
- [3] Yoshii, K. and Goto, M.: A nonparametric Bayesian multipitch analyzer based on infinite latent harmonic allocation, *IEEE Trans. on ASLP*, Vol. 20, No. 3, pp. 717–730 (2012).
- [4] Maezawa, A. et al.: Polyphonic audio-to-score alignment based on bayesian latent harmonic allocation hidden markov model, *Proc. ICASSP*, pp. 185–188 (2011).
- [5] Bello, J. P. et al.: A tutorial on onset detection in music signals, *IEEE Trans. on ASLP*, Vol. 13, No. 5, pp. 1035–1047 (2005).
- [6] Natsubayashi, K.: 初心者のためのアコースティックギターサイト, k.natsu (オンライン), 入手先 (<http://www9.ocn.ne.jp/~knatsu/chords/chords.html>) (参照 2013-8-7).
- [7] 前澤 陽ほか: 音響信号と音楽的制約を統合したバイオリンの演奏弦系列の推定, 情報処理学会研究報告.[音楽情報科学], No. 5, pp. 1–6 (2009).
- [8] Radisavljevic, A. and Driessen, P.: Path Difference Learning for Guitar Fingering Problem, *Proc. ICMC* (2004).
- [9] Tuohy, D. R. and Potter, W. D.: A genetic algorithm for the automatic generation of playable guitar tablature, *Proc. ICMC*, pp. 499–502 (2005).
- [10] 江村伯夫ほか: 弾き易さを考慮したギター・コードフォーム列探索システム, 日本音響学会誌, Vol. 64, No. 2, pp. 73–83 (2008).
- [11] Barbancho, A. M. et al.: Automatic transcription of guitar chords and fingering from audio, *IEEE Trans. on ASLP*, Vol. 20, No. 3, pp. 915–921 (2012).
- [12] Radicioni, D. and Lombardo, V.: Guitar fingering for music performance, pp. 527–530 (2005).
- [13] Goto, M. et al.: RWC music database: Popular, classical, and jazz music database, *Proc. ISMIR*, pp. 287–288 (2002).
- [14] Barbancho, I. et al.: Inharmonicity-based method for the automatic generation of guitar tablature, *IEEE Trans. on ASLP*, Vol. 20, No. 6, pp. 1857–1868 (2012).
- [15] O’Grady, P. D. and Rickard, S. T.: Automatic hexaphonic guitar transcription using non-negative constraints, *Proc. ISSC*, pp. 1–6 (2009).
- [16] Hrybyk, A. and Kim, Y.: Combined audio and video analysis for guitar chord identification, *Proc. ISMIR*, pp. 159–164 (2010).
- [17] 藤井創太ほか: Fingering Simulator: ギター単旋律の運指推定 (演奏分析・支援・加工), 情報処理学会研究報告.[音楽情報科学], Vol. 2008, No. 78, pp. 167–172 (2008).
- [18] Sakae, D. et al.: Initialization-Robust Multipitch Estimation based on Latent Harmonic Allocation using Overtone Corpus, *Proc. ICASSP*, pp. 425–428 (2012).
- [19] Emiya, V. et al.: Multipitch estimation of piano sounds using a new probabilistic spectral smoothness principle, *IEEE Trans. on ASLP*, Vol. 18, No. 6, pp. 1643–1654 (2010).
- [20] 堀 玄ほか: 入出力隠れマルコフモデルの復号によるギターのための自動編曲, 情報処理学会研究報告.[音楽情報科学], Vol. 2012, No. 3, pp. 1–6 (2012).