

ギター演奏音からの難易度調整可能なタブ譜自動生成システム

矢澤 一樹

糸山 克寿

奥乃 博

京都大学 大学院情報学研究所 知能情報学専攻

1. はじめに

音響信号からの自動採譜技術は、音楽情報処理の分野で長らく研究されてきた重要な研究課題の一つである。特にギターは演奏人口が多く、市販の楽譜の数がピアノ譜などと比較して少ないため、ギター演奏音を対象とした高性能な自動採譜技術の実現が望まれている。ギターは異なる弦で同一の音高を演奏できる楽器であり、そのためギター演奏では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。ギター用の楽譜形式の一つであるタブ譜には、押弦すべき箇所が数字によって一意に表現されており、ギター演奏者はタブ譜を用いることで運指の多様性に悩まずに直感的に演奏を行うことができる。以上のような理由から、ギター演奏者の演奏支援には音響信号からのタブ譜自動生成が有意義である。

タブ譜自動生成を演奏支援に応用するには、生成されたタブ譜がギターで演奏可能であることが不可欠である。しかし、潜在的調波配分法 (Latent Harmonic Allocation: LHA) [1] などの既存の音高推定法では、演奏楽器の種類を考慮していないために、ギターで演奏不可能な音高の組み合わせが推定されうる。例えば、7音以上の同時発音や極度に早い音高の変化などは、ギター1本では演奏することができない。このような音高推定結果を既存の運指推定法 [2, 3] に直接入力して運指を推定した場合、演奏不可能なタブ譜が出力されてしまう可能性がある。

また、ギター演奏者の演奏レベルは初心者から熟練者まで幅広く、演奏できる楽曲の難易度は個人によって大きく異なる。例えば、ギターを始めたばかりの初心者の多くは自分の好きな楽曲で練習することを望むが、それらの楽曲は演奏難易度が高いことも多く、その結果彼らの練習意欲が低下してしまうことがある。一方熟練者の場合は、多少演奏するのが難しくとも、元の楽曲をできるだけ忠実に再現することを望むと考えられる。そのため、演奏者のレベルに応じて出力タブ譜の難易度を変更できる採譜システムの実現が望ましい。

本稿では、音響信号を入力とした難易度変更可能なタブ譜自動生成手法を提案する。本手法は、LHAによって推定した誤りを含む音高推定結果を基に最適なギター運指を推定し、その後推定された最適運指を基にギターで演奏不可能な音高を排除する。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し、これを動的計画法を用いて解くことで、運指時の制約を考慮した推定を行う。さらに我々は、あらかじめタブ譜の難易度と関連の高い運指の特徴量を抽出しておき、これらの特徴量によって定義される運指コストの相対的な重要度を変更することで、出力タブ譜の難易度調整を行う。

2. 提案法の概要

2.1 難易度と関連のある運指特徴量の抽出

出力タブ譜の難易度変更を行うために、難易度に影響を与える運指特徴量の抽出を行った。初めに、先行研究

表 1: 列挙した運指特徴量および、各特徴量と難易度との相関係数。太字は、難易度との相関係数が 0.4 以上となった特徴量を示す。

	特徴量	相関係数
フォームの 押弦に関する 特徴量	手の水平位置	0.81
	手の垂直位置	-0.08
	手の回転角	-0.18
	使用指数	0.12
	隣接指間の水平幅の総和	-0.01
	隣接指間の垂直幅の総和	0.52
	隣接指間のユークリッド幅の総和	0.10
	指間の最大水平幅	0.54
	指間の最大垂直幅	-0.07
	指間の最大ユークリッド幅	0.05
	バレーの有無	0.33
	解放弦の数	-0.17
フォームの 変更に関する 特徴量	追加指数	0.52
	離弦指数	0.54
	手首の移動距離	0.61
	各指の水平移動距離の総和	0.56
	各指の垂直移動距離の総和	0.46
	各指のユークリッド移動距離の総和	0.51
	指の最大水平移動距離	0.61
指の最大垂直移動距離	0.43	
	指の最大ユークリッド移動距離	0.50

[2, 3] や筆者自身の知見を基に、難易度と関係があると思われる 21 個の運指特徴量を列挙した (表 1)。ここで、運指推定時にフォーム推移に関するマルコフ性を仮定するため、今回は時間フレーム単位で計算できる特徴量のみを調査の対象とした。次に、あらかじめ人手で 6 段階に難易度が付与された 24 個のタブ譜 [4] のそれぞれに対して、上記の特徴量の値を時間フレームごとに計算し、その後タブ譜全体での平均値を求めた。最後に、これら全てのタブ譜を用いて各特徴量の値とタブ譜の難易度との相関を求めることで、難易度変更に有用な特徴量を調べた。本稿では、相関係数が 0.4 以上となった 12 個の特徴量 a_1, \dots, a_{12} を難易度に影響を与える特徴量であると判断し、これらをタブ譜生成時に用いることとした。

2.2 音響信号からのタブ譜自動生成

上記の特徴量を基に、音響信号からのタブ譜自動生成を行う。提案法では、我々の先行研究 [5] と同様に、ギター運指を押弦フォームの時間的推移とみなし、これ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することで、運指の推定を行う。

本グラフでは、時刻 t でのフォーム c_p の使用を頂点 v_{tp} によって表し、使用フォームの変更を有向辺 $e(v_{tp}, v_{(t+1)q})$ によって表す。推定される運指の演奏可能性を保証するため、グラフの設計は以下の 3 つの制約に基づいて行う。

1. 押弦フォーム制約: 各時刻で使用されるフォーム c_p は、ギターのコード表 [6] を基にあらかじめ列挙された押弦可能フォームの中から、必ず選択される。
2. フォーム変更時刻制約: 使用フォームの変更は、あらかじめ検出されたオンセット時刻でのみ可能である。
3. 同一フォーム継続制約: フォーム変化後一定時間は、変化後のフォームを用い続けなければならない。

各有向辺 $e(v_{tp}, v_{(t+1)q})$ の重みは、以下のように楽曲再現度 (Musical Reproducibility: MR) と運指難易度 (Fingering Difficulty: FD) の 2 つの値によって定義する。

Difficulty Adjustable System of Automatic Tablature Transcription from Guitar Audio: Kazuki Yazawa, Katsutoshi Itoyama, and Hiroshi G. Okuno (Kyoto Univ.)

$$\log W(e(v_{tp}, v_{(t+1)q})) \\ = w * \log MR(X_{t+1}, c_q) + (1-w) * \log FD(c_p, c_q).$$

ここで、 w は楽曲再現度と運指難易度の相対的な重要度を調整するパラメータであり、本パラメータの大きさによって出力タブ譜の難易度を制御する。 w を小さくすることで、楽曲再現度に対する運指難易度の重要度が高まり、結果としてより運指を簡略化した初心者用のタブ譜が生成されると考えられる。逆に、 w が 1 に近くなるほど運指難易度に対する楽曲再現度の重要度が高まり、 $w = 1$ のときには運指の難易度を全く考慮せずにできるだけ元の楽曲を再現するようなタブ譜が生成される。

楽曲の再現では、メロディ、リズム、和音などの再現が重要であるが、本稿では楽曲再現度として特に和音再現度 (Pitch Reproducibility: PR) を考えることとする。和音再現度は、既存の音高推定法 LHA [1] の推定結果に基づいて定義する。LHA では、楽器音のスペクトログラム X を入力として与えることで、各時刻 t における各音高 k の出現度 N_{tk} を求めることができる。フォーム c_p による時刻 t での和音再現度を、本フォームで演奏可能な 6 つの音高 $k_{c_{p1}}, \dots, k_{c_{p6}}$ の出現度の総和によって定義する。

$$MR(X_t, c_p) \approx PR(N_t, c_p) \propto \sum_{k=k_{c_{p1}}}^{k_{c_{p6}}} N_{tk}.$$

また運指難易度は、難易度と関連があると判断された 12 個の特徴量の線形和によって運指コストを計算し、その逆数によって定義する。

$$FD(c_p, c_q) \propto \frac{1}{\sum_{i=1}^{12} \theta_i a_i(c_p, c_q)}.$$

ここで、 θ_i は各特徴量 a_i の相対的な重みを決定するパラメータである。

こうして設計された重み付き有向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、最適フォーム系列 c_1^*, \dots, c_T^* を推定する。最後に、各時刻で推定された最適フォーム c_i^* で演奏可能な全ての音高について、その出現度を閾値判定することによって、音高の有無を判断する。すなわち、音高 k がフォーム c_i^* で演奏可能であり、さらに N_{ik} が特定の閾値 $\alpha * \max_{ik} N_{ik}$ 以上である場合のみ、その音高が時刻 t で鳴っているとみなす。推定された最適フォーム系列 c_1^*, \dots, c_T^* および音高出現度の閾値判定結果を用いて、最終的なタブ譜を生成する。より厳密には、実際のタブ譜生成にはビート検出やテンポ推定などが別途必要となるが、本稿では紙面の都合上省略する。

3. 評価実験

本システムの有効性を確認するため、音高再現度と運指難易度の相対重みパラメータ w を変化させたときの、出力タブ譜の音高推定精度と運指難易度の比較を行った。

3.1 実験条件

実験データには、RWC 研究用音楽データベース [7] から抽出した計 93 のギターソロパートの、各冒頭 60 秒を用いた。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。

音高推定精度の評価基準には、時間フレーム単位での F0 の再現率・適合率・F 値を用いた。また運指難易度の評価には、本稿で定めた運指コストのパート全体での合計値および、パート中に登場するフォームの総数を用いた。

表 2: 提案法および LHA の音高・運指推定結果。値は全パートに対する平均値を表す。

w	LHA	提案法					
		1.00	0.90	0.80	0.70	0.60	0.50
F 値	0.69	0.73	0.74	0.70	0.61	0.51	0.43
適合率	0.70	0.74	0.75	0.76	0.75	0.72	0.68
再現率	0.70	0.73	0.74	0.66	0.52	0.41	0.32
出現フォーム数	-	33.6	25.7	19.3	11.8	6.9	4.4
運指コスト	-	1.83	1.17	0.55	0.25	0.12	0.07

本稿では、パラメータ w の値を 1.0 から 0.5 まで 0.1 刻みで変化させ、それぞれの条件で推定される音高・運指の比較を行った。運指コスト計算時の各特徴量の相対重み θ については、今回は手動で設定した。またシステムの最大性能を比較するため、音高推定の閾値パラメータ α は各パート・条件ごとに F 値が最大となるように最適化した。同一フォーム最低継続時間は、一般的なギター演奏者が 1 秒間にフォームを 5 回以上変更するのは不可能であると考え、200ms とした。音高推定精度の比較のため、同一の実験データを用いて従来法である LHA による音高推定も行った。

3.2 実験結果

提案法と LHA による音高・運指推定結果を、表 2 に示す。表より、 w の値が大きいたまには、提案法の音高推定精度が F 値・適合率・再現率ともに LHA の値を上回っていることが分かる。これは、運指推定時に用いた 3 つの制約によって、LHA の推定結果に含まれる不適切な音高が効果的に排除できていることを示す。

w の値を小さくしたときには、音高の再現率は低下するものの、適合率は LHA の値と同程度に保たれた。またこのとき、出現フォーム数、運指コストがともに減少した。このことより、 w を小さくして初心者用のタブ譜を生成した場合、音の取りこぼしが増えるものの、余分な音高を推定せずに運指を簡略化できることが分かった。

4. おわりに

本稿では、タブ譜の難易度に影響を与えるギター運指の特徴量をあらかじめ抽出し、得られた特徴量を基にグラフ探索法によって最適な運指を決定することで、運指可能性と演奏難易度を考慮したタブ譜生成を行った。今後は、難易度変更用いる運指特徴量の追加や、決定木や重回帰分析による判別ルールの学習などを行うことで、より厳密な難易度調整を実現したい。また、音高・運指推定のアルゴリズムを工夫することで、初心者用のタブ譜を生成する際にメロディラインなどの重要な音高のみを優先して残すことに取り組みたい。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] 吉井和佳, 後藤真孝: “多重基本周波数解析のための無限潜在的調波配分法”, 情報処理学会 音楽情報科学研究会 研究報告 MUS-86, Vol.2010, No.11, pp.1-9, 2010.
- [2] A. Radisavljevic and P. Driessen, “Path Difference Learning for Guitar Fingering Problem”, *Proc. of ICMC*, Vol.28, 2004.
- [3] D. Radicioni and V. Lombardo, “Guitar Fingering for Music Performance”, *strings*, Vol.40, No.45, p.50, 2005.
- [4] k. yamane: “無料タブ譜なら PTB ファイル・ギター TAB 譜 (タブ譜)”, <http://ptb.jp/>, (参照 2014-1-10).
- [5] 矢澤一樹, 糸山克寿, 奥乃博: “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”, 情報処理学会 音楽情報科学研究会 研究報告 MUS-100, Vol.2013, No.17, pp.1-6, 2013.
- [6] mamoo: “mamoo”, <http://www.geocities.jp/mamoo/>, (参照 2014-1-10).
- [7] 後藤真孝, 橋口博樹, 西村拓一, 岡隆一: “RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース”, 情報処理学会論文誌, Vol.45, No.3, pp.728-738, 2004.