

押弦制約と運指制約を用いたタブ譜自動生成システム

矢澤 一樹^{1,a)} 阪上 大地^{1,b)} 糸山 克寿^{1,c)} 奥乃 博^{1,d)}

概要:

本稿は、ギター演奏者の押弦・運指に関する制約を用いたタブ譜自動生成システムについて報告する。我々は、潜在的調波配分法 (LHA) の推定結果に押弦・運指制約を加えることで、演奏不可能な音の組合せを排除する。押弦制約として、あらかじめ列挙した演奏可能な押弦パターンの中から各時間フレームでの最適パターンを推定し、その押弦パターンで演奏可能な音のみを抽出する。これにより、演奏可能な押弦パターンのみで構成されたタブ譜を出力することができる。また、運指制約として編集距離に基づく運指のコストを用いて、運指の面でもタブ譜の改善を行った。実験の結果、我々はさらに本手法が音高推定を頑健に行えることも確認した。

Automatic Guitar Tablature Production System based on Configuration and Fingering Constraints

Abstract: This paper describes an automatic tablature generation system for guitar performances using constraints of finger configurations and fingering. We exclude unplayable combinations of sounds from estimation results of the latent harmonic allocation (LHA) by using the constraints. Our system estimates optimal finger configurations for each time frame among playable finger configurations and extract playable sounds on the configurations. Thus our system can output a tablature consisting of only playable finger configurations. In addition, we define the fingering constraint as a cost function based on edit distance to improve the quality of generated tabulatures. Experimental results showed that our system achieved robust multipitch estimation by using the constraints.

1. はじめに

ギターは最もポピュラーな楽器の一つであり、初心者から熟練者まで演奏者の層は幅広い。従来から、ギター演奏者の演奏支援技術として楽譜追跡 [11] やタブ譜検索 [10] が取り組まれている。ギター演奏熟練者は五線譜 (どの音を演奏するかが記述されている) があれば楽曲を演奏できるが、ギター演奏初心者にはそれは難しく、多くの場合タブ譜 (どの弦のどの位置を押弦するかが記述されている) を演奏に用いる。しかし、五線譜の数に比べてタブ譜の数は不十分であり、演奏したい楽曲のタブ譜を入手することは必ずしも容易ではない。したがって、ギター演奏初心者の演奏支援にはタブ譜の自動生成が有益である。

音楽音響信号から様々な楽譜を自動生成する自動採譜は

1970年代から取り組まれている研究分野であり、自動タブ譜推定との関連も深い。自動採譜における最も重要な課題の一つが複数同時発音からそれぞれの音の高さを推定する多重基本周波数推定であり、ギターの和音解析などにも応用できる。多重基本周波数推定技術として、人間の聴覚機構をモデル化し計算機上に実現したもの [9] などが取り組まれており、近年では音楽音響信号の時間周波数表現を確率分布とみなし、混合正規分布 (Gaussian mixture model; GMM) で表現される音モデルを観測分布に対して適応させる手法 [5, 8, 15] がよく用いられている。潜在的調波配分法 (latent harmonic allocation; LHA) [15] は GMM を用いた高性能な多重基本周波数推定法であるが、LHA の結果をそのままタブ譜生成に用いることはできない。問題の一つは、楽器の構造や身体の制約上人間が演奏できないような音の組合せを LHA が推定することである。例えば通常のギターは 6 弦であるが、LHA は 7 音以上の同時発音を推定することがあり、このような推定結果はタブ譜に直接変換することができない。また 6 音以下の音の組合せであっ

¹ 京都大学 大学院情報学専攻 知能情報学専攻

a) kyazawa@kuis.kyoto-u.ac.jp

b) dsakaue@kuis.kyoto-u.ac.jp

c) itoyama@kuis.kyoto-u.ac.jp

d) okuno@kuis.kyoto-u.ac.jp

ても、人間が同時には演奏できないような組合せが推定される場合があり、そのような結果をタブ譜に変換したとしても、演奏支援に用いることはできない。

これらの問題の解決には、演奏楽器をピアノやギターなど特定の楽器に限定することで、楽器の音色、音域、演奏可能な音の組合せなどを限定することが有効である。例えば、Emiya はピアノの調波構造の性質をモデルに取り入れることで、ピアノ演奏に特化した高性能な多重基本周波数推定法を提案している [4]。Barbancho らは、隠れマルコフモデル (hidden Markov model; HMM) を用いてギター演奏音からコードと押弦パターンを推定し、自動採譜を行う方法を提案した [1]。彼らは 330 個の押弦パターンを演奏可能として定めているが、このパターン数は不十分であると考えられる。なぜなら、特にジャズなどのギター演奏では、それら 330 個のパターンに含まれない特殊なパターンが頻繁に用いられるからである。そのようなパターンはギター演奏で必ずしも一般的に用いられるものではないが、実際に人間が演奏することができるため、そのようなパターンの存在を無視すべきではない。したがって、押弦可能パターンを演奏者の身体的制約上演奏できるか否かという基準だけで選択することが望ましい。

本稿では、LHA に押弦・運指制約を加えることで、ギターで演奏可能な音の組合せのみを得る手法を報告する (図 1)。本手法では、演奏可能な押弦パターンをあらかじめ列挙し、その後各時間フレームでの最適パターンを決定することで、人間が演奏不可能な音の組合せを抑圧する。押弦可能パターンはフレット幅と押弦箇所数の 2 条件で選別し、またギター演奏で頻繁に用いられるセーハと呼ばれる演奏法もパターンに含める。さらに、運指に関する制約を同時に加えることで、より人間が演奏しやすい押弦パターン系列を推定する。実験では、本制約によって演奏可能なタブ譜のみを出力可能になったことに加え、音高の推定精度に関する頑健性も同時に得られたことを示す。

2. 潜在的調波配分法の概要

本稿では潜在的調波配分法 (LHA) [15] に押弦制約と運指制約を組み込んだ手法について述べる。そのため、本節では LHA の概要について述べる。

LHA でははじめに、観測された音響信号を連続ウェーブレット変換し、得られた振幅スペクトルの系列に対して基本周波数推定を行う。時間フレーム数を D とし、 D フレーム合わせた全ての観測変数を $X = \{X_1, \dots, X_D\}$ とする。ここで $X_d = \{x_{d1}, \dots, x_{dN_d}\}$ は各フレームにおける観測周波数の系列であり、例えばフレーム d の振幅スペクトルにおける周波数 f のパワーが a であれば、フレーム d で周波数 f は a 回観測されたとみなす。 N_d はフレーム d での観測周波数の総数である。

次に上記の振幅スペクトルを、それぞれが倍音数 M の

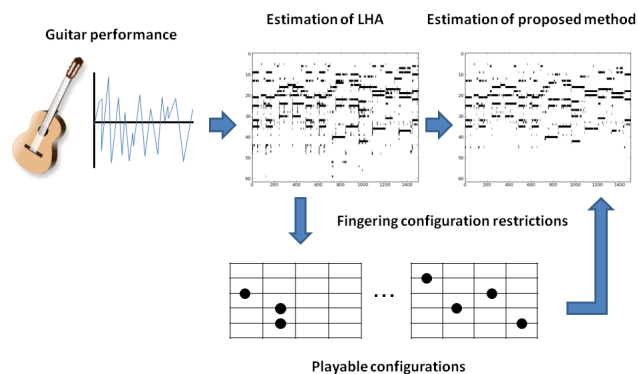


図 1 本手法の流れ。Estimation of LHA と Estimation of proposed method はそれぞれ従来法、本手法によって生成されたピアノロールを表わしている。

調波構造をもつ K 個の基底を混合した、以下のネスト型混合ガウス分布で定式化する。

$$\mathcal{M}_k(x) = \sum_{m=1}^M \tau_{km} \mathcal{N}(x | \mu_k + o_m, \lambda_k^{-1})$$

$$\mathcal{M}_d(x) = \sum_{k=1}^K \pi_{dk} \mathcal{M}_k(x)$$

$\mathcal{N}(x | \mu, \lambda^{-1})$ は平均が μ 、精度 (分散の逆数) が λ の正規分布の確率密度関数を表わす。また μ_k 、 λ_k は基底 k のガウス分布の平均および精度であり、 o_m は M 個のガウス分布を倍音関係に配置するオフセット値である。さらに τ_{km} は基底 k における倍音 m の相対強度、 π_{dk} はフレーム d での基底 k の相対強度である。

観測変数 X に対する潜在変数として、 $Z = \{Z_1, \dots, Z_D\}$ を導入する。ここで、 $Z_d = \{z_{d1}, \dots, z_{dN_d}\}$ である。 z_{dn} は $K \times M$ 次元のベクトルであり、 x_{dn} が k 番目の基底の m 番目の倍音から生成されたとき、 $z_{dnkm} = 1$ とする。観測モデルは以下ようになる。

$$p(X|Z, \mu, \lambda) = \prod_{dnkm} \mathcal{N}(x_{dn} | \mu_k + o_m, \lambda_k^{-1})^{z_{dnkm}}$$

$$p(Z|\pi, \tau) = \prod_{dnkm} (\pi_{dk} \tau_{km})^{z_{dnkm}}$$

$$p(\pi) = \prod_{d=1}^D \text{Dir}(\pi_d | \alpha_0) \propto \prod_{d=1}^D \prod_{k=1}^K \pi_{dk}^{\alpha_0 - 1}$$

$$p(\tau) = \prod_{k=1}^K \text{Dir}(\tau_k | \beta_0) \propto \prod_{k=1}^K \prod_{m=1}^M \tau_{km}^{\beta_0 - 1}$$

$$p(\mu, \lambda) = \prod_{k=1}^K \mathcal{N}(\mu_k | m_0, (\gamma_0 \lambda_k)^{-1}) \mathcal{W}(\lambda_k | w_0, \delta_0)$$

$\text{Dir}(\cdot)$ はディリクレ分布の確率密度関数、 $\mathcal{W}(\cdot)$ はウィシャート分布の確率密度関数、 $\alpha_0, \beta_0, m_0, \gamma_0, w_0, \delta_0$ はモデルのハイパーパラメータである。

潜在変数の事後分布を

$$p(Z, \pi, \tau, \mu, \lambda | X) \approx q(Z)q(\pi, \tau, \mu, \lambda)$$

と因子分解する．各パラメータは変分 EM アルゴリズムで推定する．

変分 E ステップでは， π, τ, μ, λ の現在の値を使用し，クラス割り当て z_{dnkm} の期待値 ρ_{dnkm} を更新する．

$$\begin{aligned} \log q^*(Z) &= \mathbb{E}_{\pi, \tau, \mu, \lambda}[\log p(X, Z, \pi, \tau, \mu, \lambda)] + \text{const.} \\ &= \sum_{dnkm} z_{dnkm} \log \rho_{dnkm} + \text{const.} \\ \log \rho_{dnkm} &= \mathbb{E}[\log \pi_{dk}] + \mathbb{E}[\log \tau_{km}] \\ &\quad + \mathbb{E}[\log \mathcal{N}(x_{dn} | \mu_k + o_m, \lambda_k^{-1})] \end{aligned}$$

変分 M ステップでは，十分統計量 N_k, N_{dk}, N_{fkm} の値を計算し， π, τ, μ, λ の変分事後分布を更新する．共役事前分布を使用しているため，変数の事後分布は全て事前分布と同じ形になる．

$$\begin{aligned} N_k &= \sum_{dnm} \rho_{dnkm} \\ N_{dk} &= \sum_{nm} \rho_{dnkm} \\ N_{fkm} &= \sum_d \sum_{x_{dn}=x_f} \rho_{dnkm} \\ q(\pi, \tau, \mu, \lambda) &= \prod_{d=1}^D q(\pi_d) \prod_{k=1}^K \{q(\tau_k)q(\mu_k, \lambda_k)\} \\ q(\pi_d) &= \text{Dir}(\pi_d | \alpha_d) \\ q(\tau_k) &= \text{Dir}(\tau_k | \beta_k) \\ q(\mu_k, \lambda_k) &= \mathcal{N}(\mu_k | m_k, (\gamma_k \lambda_k)^{-1}) \mathcal{W}(\lambda_k | w_k, \delta_k) \end{aligned}$$

事後ハイパーパラメータ $\alpha_{dk}, \beta_{km}, \gamma_k, \delta_k, m_k, w_k$ を計算する．

$$\alpha_{dk} = \alpha_0 + N_{dk}, \quad \beta_{km} = \beta_0 + N_{km} \quad (1)$$

$$\gamma_k = \gamma_0 + N_k, \quad \delta_k = \delta_0 + N_k \quad (2)$$

$$m_k = \frac{\gamma_0 m_0 + \sum_{fm} N_{fkm} (x_f - o_m)}{\gamma_0 + N_k} \quad (3)$$

$$w_k^{-1} = w_0^{-1} + \gamma_0 m_0^2 + \sum_{fm} N_{fkm} (x_f - o_m)^2 - \gamma_k m_k^2 \quad (4)$$

ここで， x_f は f 番目の周波数ピンの対数周波数である．

パラメータの推定後，フレーム d における基底 k の有効観測数 $N_{dk} = \pi_{dk} N_d$ がある閾値以上となるような k を求め，その音高の音を実際に演奏された音と判断する．すなわち t を閾値パラメータとし， $N_{dk} \geq t \times \max_{dk} N_{dk}$ を満たす全ての k に対応する音高をフレーム d での推定結果とする．ここで，後述する押弦制約のため， N_{dk} が閾値未満であるような音高 k については $N_{dk} = 0$ としておく．

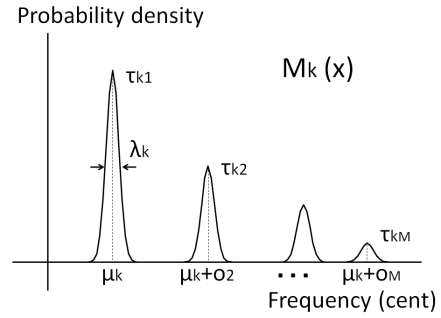


図 2 調波構造モデル $M_k(x)$. これは M 個のガウス分布の組合せで表現されている． μ_k はガウス分布の位置， λ_k は分布の精度， o_m はオフセット， τ_{km} は m 番目の倍音成分の重みを表す．

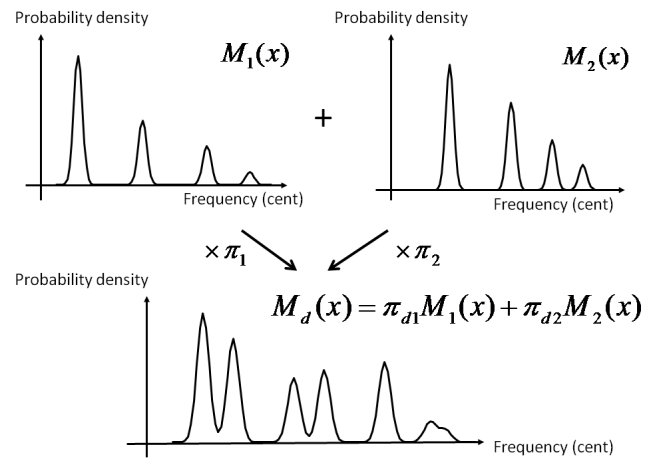


図 3 ネスト型混合ガウス分布 $M_d(x)$. M_d は， K 個の調波構造の π_k による重み付け和として得られる．(図は $K = 2$ の場合．)

3. 押弦制約

我々は押弦パターンを押弦制約として用いることで，ギターのタブ譜を出力可能な多重基本周波数推定法を構築する．押弦パターンとは，タブ譜の一行分に対応づけられる，ギターの各弦の押弦位置の組合せを表す 6 つの数の組のうち，人間が演奏可能なものである．押弦パターンの演奏可能性は押弦箇所数（同時に押さえる弦の数）とフレット幅（指を広げる幅）という 2 つの条件に基づいて判断する．ギター演奏においては基本的には 1 本の指で 1 本の弦を押さえるが，人差し指で複数の弦を同時に押さえるセーハ（図 4）も頻りに演奏に用いられるため，セーハを利用したパターンも押弦パターンに含める．

3.1 押弦パターンの列挙

はじめに，手首の位置を考えず指の相対的な位置関係だけを考えた基本パターンを生成する．後でセーハを考えるため，人差し指を使用せずに押弦可能な基本パターン (BP1) と人差し指の使用を許した基本パターン (BP2) を，それぞれの条件で列挙する．

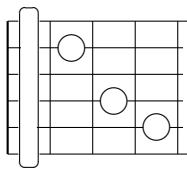


図 4 セーハの例．長方形は人差し指に対応．

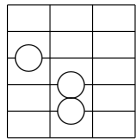


図 5 BP1 の例.

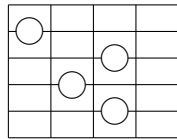


図 6 BP2 の例.

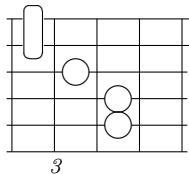


図 7 P1 の例．数字 3 は
 フレット番号を表す．

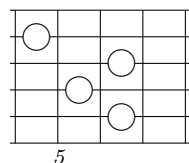


図 8 P2 の例.

BP1 押弦箇所数 3 以下，フレット幅 3 以下．

BP2 押弦箇所数 4 以下，フレット幅 4 以下．

次に BP1 と BP2 をギター の指板上で水平方向に移動することで，手首の位置を考慮した押弦可能パターン P1, P2 を作成する．標準的なギターは 21 のフレットをもつため，BP1 の一つのパターンの移動で 19 通り，BP2 の一つのパターンの移動では 18 通りのパターンが生成できる．さらに，セーハを用いた押弦パターンを P1 に追加する．これは，BP1 の左隣のフレット上にある最大 6 本の弦を人差し指で押さえるものである．

P1 BP1 を水平方向に移動し (19 通り)，左隣の複数の弦 (0-6 本) をセーハしたパターン (図 5) ．

P2 BP2 を水平方向に移動 (18 通り) したパターン ．

最後に，P1 と P2 を合わせ重複したパターンを削除した全てのパターンを，押弦可能パターン C_1, \dots, C_P とする．本手法で列挙された押弦可能パターンの総数 P は 89479 となった．各パターンで発音可能な音の組合せを K_p とする．ここでは通常の 6 弦ギターを考えているので，各組合せ K_p には 6 つの音が含まれる (ただし重複を許す) ．

3.2 押弦制約の適用

LHA を用いたタブ譜推定における問題は，楽器の構造や身体的制約を考慮しないため，ギター演奏として不自然な音の組合せが生成されうることである．そこで本手法では LHA による推定結果を後処理で修正し，時間フレームごとに最適押弦パターンを推定してそのパターンで発音不可能な音を排除する ．

事前に列挙した押弦可能なパターン C_1, \dots, C_P ごとにフレーム d でのパターン C_p の有効観測数 $N_{dp} = \sum_{k \in K_p} N_{dk}$ を計算する．そして，各フレーム d でこの値が最大となる

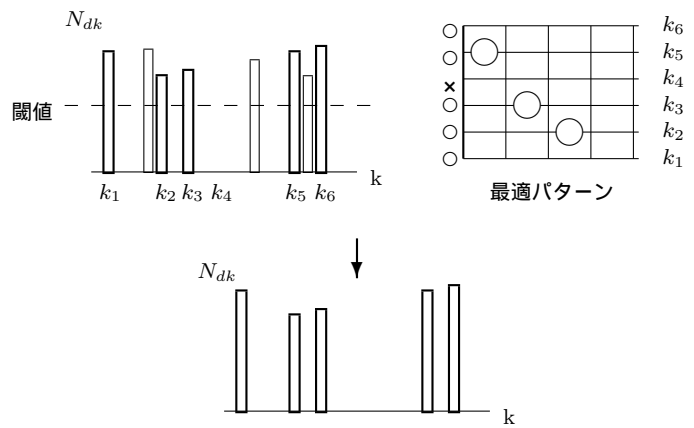


図 9 押弦制約適用のイメージ．ここで最適パターンとは， N_{dp} を最大化する押弦パターン C_p である． N_{dp} は N_{dk} の対応するノートナンバーの値の和として得られる ．

ような p に対応するパターン C_p をそのフレームでの最適パターンとする．ただし後述する運指制約のため，ここでは N_{dp} の値が最大になるようなパターン C_{p_d} (第 1 最適パターン) の他に，2 番目に値が大きくなるようなパターン C_{q_d} (第 2 最適パターン)，3 番目に値が大きくなるようなパターン C_{r_d} を同時に求めておく．これらを組み合わせた $\{C_{p_d}, C_{q_d}, C_{r_d}\}$ を時刻 d での最適フレーム候補とし，次節の運指制約によって最終的な最適パターンを各時刻ごとに一つ決定する ．

4. 運指制約

前節までの手法は，時間フレームごとに独立した観測に対して最適押弦パターンと多重基本周波数を推定するものであった．しかし現実的には，隣接する時間フレームの間で押弦パターンが独立に定まることはほとんどない．ある音を発音している間は，手や指のフォームは一定と考えられる．すなわち楽曲全体として見た場合，押弦パターンは隣接時刻間である程度同じところに留まろうとする ．

また，押弦パターン決定の際には，以下の点についても考慮する必要がある．ギターは異なる弦で同じ高さの音を出すことができる楽器であり，そのためある音の組合せを複数の異なる押弦パターンで演奏することができる．例えば，パワーコード B5 (B1, F#2, B2 の音の組合せ．) は 2 つの異なった押弦パターンを用いて演奏することができる (図 10) ．このことは，我々が音高の組合せだけから押弦パターンを一意に決定できないということを意味する．実際，押弦制約によって求めた時刻 d での最適パターン C_{p_d} は，その構成音が 5 音以下である場合，第 2 最適パターン C_{q_d} や第 3 最適パターン C_{r_d} と全く同じ構成音を持つこともある．ギターの演奏者は一般的に，複数ある押弦パターン候補の中から運指の容易なパターンを選んで演奏を行う ．

したがって本節ではこれらの知見を考慮して，最終的な最適パターン推定を行う．すなわち，押弦制約によって得

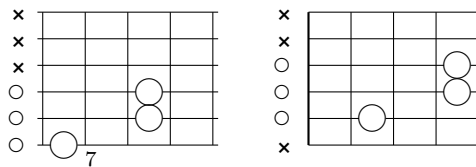


図 10 異なる押弦パターンで同一の音高の組合せ (B1, F#2, B2) を生じる例。図の左の および×は、それぞれ対応する弦を弾く、弾かないを表わしている。

られた最適パターン候補に対して、さらに運指に関する制約を加えることで、より人間が演奏しやすい押弦パターン系列を一つ推定する。このような制約を、我々は運指制約と呼ぶこととする。

4.1 運指制約の適用

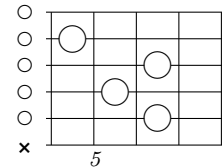
押弦パターンを、第 1 弦から第 6 弦までの押弦位置を表す 6 つの数の組で表現する (図 11)。このように表現される押弦パターンの移動コストを押弦パターン間の編集距離で定義する。現在のフレーム d の押弦パターンを $C_d = (a_{d1}, \dots, a_{d6})$, その直前のフレーム $d-1$ での押弦パターンを $C_{d-1} = (a_{d-1,1}, \dots, a_{d-1,6})$ とし、2 つのパターン間での運指コスト $FC(P_d, P_{d-1})$ を以下のように定める。

- $C_d = C_{d-1}$, すなわちパターンの変化が無い場合。
 $FC(C_d, C_{d-1}) = 0$.
- $\forall j, a_{d,j} = a_{d-1,j} + i, (-21 \leq i \leq 21)$, すなわち、指の配置が変わらず手首の移動だけでパターンの変更が行える場合。
 $FC(C_d, C_{d-1}) = 1$.
- 上記以外の場合。
 $FC(C_d, C_{d-1})$
 $= \sum_{\{j \text{ s.t. } (a_{d,j} \neq -) \cap (a_{d-1,j} \neq -)\}} |a_{d,j} - a_{d-1,j}|$
 (ただし、 $a_{d,j} = 0$ または $a_{d-1,j} = 0$ のときには、 $|a_{d,j} - a_{d-1,j}| = 1$ とする) .

上記の運指コストに基づいて、全体のコスト $\sum_{d=1}^{D-1} FC(C_d, C_{d-1})$ が最小となるように、前述した最適パターン候補の中から最適パターンを各時刻ごとに 1 つずつ決定する。これにより、より運指の容易な押弦パターン系列を推定することができる。

4.2 発音区間による多数決処理

上記の運指制約によって、運指コストの少ないタブ譜を出力することが可能になったが、図 12 のような運指コストの少ないわずかなパターン変化が頻繁に発生するという問題がまだ残っている。このような押弦パターンの揺らぎを少なくするため、発音区間による多数決処理を行う。発音区間の境界は、音響信号をウェーブレット変換後、spectral flux に基づいて推定する。spectral flux は隣接時刻間での周波数スペクトルのパワー変化を表した尺度で



(0, 5, 7, 6, 7, -)

図 11 押弦パターンの表現例。第 1 弦 (図の一番上の弦) から第 6 弦 (図の一番下の弦) までの押弦箇所のフレット番号で押弦パターンを表す。ここで 0 は解放弦 (弦を左手で押さえずに鳴らすこと) を表し、"- " はミュート (弦を鳴らさないこと) を表す。

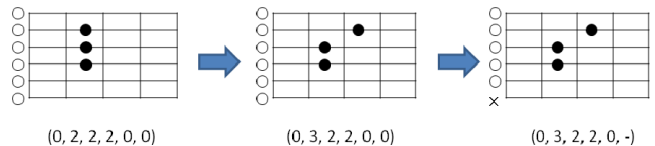


図 12 運指コストの少ない押弦パターン推移の例。このようなパターン変化が短時間で過度に頻繁に起こる場合、推定誤りであると考えられる。

あり、以下のように定義される。

$$SF_d = \sum_{\{f \text{ s.t. } SP_{d,f} > SP_{d-1,f}\}} (SP_{d,f} - SP_{d-1,f})$$

f は周波数の高さを表し、 $SP_{d,f}$ は時刻 d での周波数 f のパワーを表す。

SF_d を各時刻で計算し、この値が適当な閾値以上となるような全ての時刻を発音区間の境界とする。次にそれぞれの発音区間で、最もよく出現する押弦パターンをその区間での最頻パターンと定める。最後に、その区間での押弦パターンおよび音高の推定結果を、この最頻パターンおよび対応する音高に置き換える。

5. 評価実験

本手法の性能を評価するため、LHA に押弦制約を加えた本システムと制約を加えていない従来のシステムの両方を使って、ギター演奏音から時間フレームごとの F_0 を推定する実験を行い、結果を比較した。

5.1 実験条件

実験データには、RWC 研究用音楽データベース [5] に含まれるジャズ楽曲のうち、ギター 1 本で演奏された 4 曲の冒頭 24 秒を用いた (ここで、RM-J008 はギターのみを用いた楽曲であるが、2 本のギターで演奏されていると考えられるため、今回は用いていない)。楽曲は全て MIDI 音源 (YAMAHA MOTIF-XS) を使用して音響信号を作成し、ウェーブレット変換を行った。正解データ (各時間フレームでの F_0 の正解値) は MIDI ファイルから作成し、 F_0 推定の評価尺度として時間フレームレベルでの F 値を用いた。モデル中の各パラメータは変分 EM アルゴリズムを 100

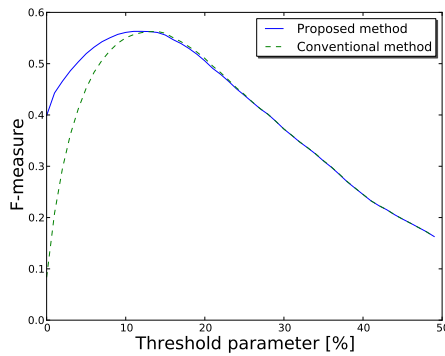


図 13 スレッシュドごとの F 値の変化．横軸はスレッシュド，縦軸は 4 曲の F 値の平均を表す．

回繰り返すことで推定し，事前分布は全て無情報事前分布とした．基底数 K および倍音数 M はそれぞれ 73, 6 とし，閾値パラメータ t については各楽曲に対する最適値，0.20, 0.10, 0.05, 0.01 の 5 条件で実験を行った．また従来法と性能を比較するため，上記と同様の条件で押弦制約を用いない通常の LHA による推定も行った．

5.2 実験結果

表 1 に実験結果を示す．最適な閾値を用いたときの F 値は従来法と提案法で等しいが，提案法では閾値を変化させたときに F 値が低下しにくいことがわかる．すなわち，提案法の閾値に対する頑健性が示されている．

6. 考察

6.1 閾値変化に対する頑健性

本システムは従来の手法に比べ，閾値の変化に対してより頑健であった．このことは，既存の LHA で閾値を下げたときに推定されてしまう余分な（実際には演奏されていない）音が，押弦制約によって抑圧されていることを示している．このような閾値頑健性のおかげで，ユーザーは演奏環境や楽曲ごとに閾値を調整しなくても安定した認識精度を得ることができる．

6.2 タブ譜の生成

本システムのもう一つの重要な利点は，推定された各時間フレームでの最適押弦パターンを用いることで，ギタータブ譜を生成することができる点である．出力されたタブ譜の例を，タブ譜に現れている押弦パターンのイラストとともに，図 14 に載せた．最適パターンは事前に列挙された押弦可能パターン C_1, \dots, C_P から選択されたものであるため，これらのタブ譜には演奏可能なパターンしか含まれない．

また，ギター演奏では異なる弦を用いて同じ高さの音を出すことができるため，音高の組み合わせだけから押弦パターンを一意に定められないということはすでに述べた．

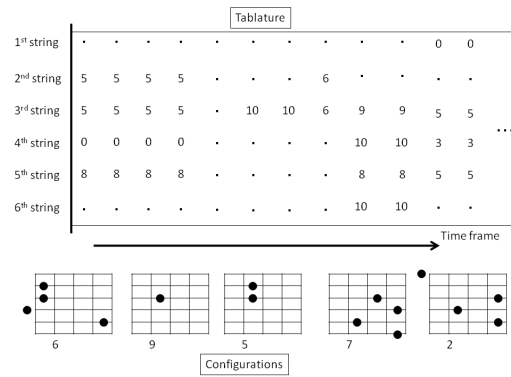


図 14 本システムで生成したタブ譜の例．RM-J009 の最初の 180 フレーム（約 3 秒）に対応している．下の図は押弦パターンの変化（運指）を表す．

本稿ではこの問題を解決するため，各時刻での最適押弦パターンを第三候補まで考慮し，これに運指制約を加えてより運指の容易なパターン系列を推定した．これにより，音高の推定精度を損なわずに，より運指に適したタブ譜を一意に生成することができる．

ただしこのタブ譜は，各時刻での音高の組み合わせ，押弦の可能性，運指の容易さによって決定したものであり，実際に演奏された運指と同じ運指が用いられているという保証はない．正確な（実際に演奏された）運指を推定するための方法としては，視覚情報を用いたもの [3, 7, 13] や特別なマイクをギターに取り付ける方法 [12] などがある．また，ソロギターのような楽曲ではなくギターの伴奏曲などが対象である場合には，コードの推移傾向などを用いてギターのコードをあらかじめ推定し，その後それらのコードに対して主要な押弦パターンを一意に割り当てるというのも有効な方法であると考えられる．

6.3 押弦可能パターンの妥当性

列挙した押弦可能パターンは，主要なギターコード表 [16] に載っている押弦パターンを全て含んでいるため，一般的に用いられるパターンの大部分をカバーしていると考えられる．しかし，パターンの過不足の問題は依然として残っている．すなわち，我々の押弦可能パターンに含まれない特殊なパターン（親指を使ったもの，中指や薬指でセーハを行ったもの，など）も存在し，また逆に実際には演奏不可能であったりほとんど用いられないようなパターンが含まれてしまっているという問題もある．このことは，我々のパターン列挙法には改善の余地がかなりあることを示している．他のパターン列挙法としては，ギターのコード表からパターンを抽出する方法や，人間の指の構造をより厳密に考える方法，先にコードを推定し後から押弦パターンを当てはめる方法などが考えられる．

表 1 基本周波数推定結果の F 値．最適値とは各楽曲で F 値がもっとも大きくなるような閾値パラメータの値を意味し，括弧内の数字は各楽曲に対する実際の値を表わす．

閾値変数	LHA(従来法)					押弦制約付き LHA(提案法)				
	最適値 (閾値)	0.20	0.10	0.05	0.01	最適値 (閾値)	0.20	0.10	0.05	0.01
RM-J006	0.64(0.17)	0.62	0.56	0.40	0.14	0.64(0.16)	0.62	0.60	0.51	0.40
RM-J007	0.83(0.09)	0.59	0.83	0.73	0.35	0.83(0.08)	0.54	0.83	0.81	0.73
RM-J009	0.87(0.15)	0.84	0.73	0.56	0.24	0.87(0.15)	0.86	0.81	0.71	0.59
RM-J010	0.83(0.17)	0.82	0.72	0.58	0.26	0.86(0.17)	0.83	0.79	0.75	0.66
平均	0.79	0.72	0.71	0.57	0.25	0.79	0.71	0.75	0.70	0.59

7. おわりに

本稿では，多重基本周波数推定法 LHA にギターの押弦制約および運指制約を加える方法について述べた．本システムは，押弦・運指制約によって求めた各時間フレームでの最適パターンを利用することで，人間が演奏するのに適したタブ譜を出力できる．またもう 1 つの利点として，押弦制約による閾値頑健性のため，ユーザーが閾値を演奏環境や楽曲に応じて調整しなくても安定した認識精度が得られる．さらに本手法では演奏楽器をギターに限定していたが，押弦パターンの列挙法を変えることで他の弦楽器の採譜にも応用することが可能である．

ただ，本手法で列挙した押弦可能パターンには過不足の問題が残されており，列挙法にはまだ改善の余地がある．今後，その改善に取り組みたい．また，HMM などを用いて運指制約をより厳密に定める，音楽的なコードの推移条件を用いる，などの改善にも取り組む予定である．

謝辞 本研究の一部は科研費 (S) No. 24220006 の支援を受けた．

参考文献

- [1] A. M. Barbancho, A. Klapuri, L. J. Tardon, and I. Barbancho: "Automatic transcription of guitar chords and fingering from audio," *IEEE Trans. on ASLP*, Vol. 20, No. 3, pp. 915–921, 2012.
- [2] I. Barbancho, L. J. Tardon, A. M. Barbancho, and S. Sammartino: "Pitch and played string estimation in classical and acoustic guitars," *AES*, pp. 1–9, 2009.
- [3] A. Burns and M. Wanderley: "Visual methods for the retrieval of guitarist fingering," *NIME*, pp. 196–100, 2006.
- [4] V. Emiya, R. Badeau, and B. David: "Multipitch estimation of piano sounds using a new probabilistic spectral smoothness principle," *IEEE Trans. on ASLP*, Vol. 18, No. 6, pp. 1643–1654, 2010.
- [5] M. Goto: "A real-time music scene description system: Predominant-F0 estimation for detecting melody and bass lines in real-world audio signals," *Speech Commun.*, Vol. 43, no. 4, pp. 311–329, 2004.
- [6] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, "RWC music database: Popular, classical, and jazz music database," *ISMIR*, pp. 287–288, 2002.
- [7] A. Hrybyk and Y. Kim: "Combined audio and video analysis for guitar chord identification," *ISMIR*, pp. 159–164, 2010.
- [8] H. Kameoka, T. Nishimoto, and D. Barber: "A multi-

- pitch analyzer based on harmonic temporal structured clustering," *IEEE Trans. on ASLP*, Vol. 15, no. 3, pp. 982–994, 2007.
- [9] A. Klapuri: "Multipitch analysis of polyphonic music and speech signals using an auditory model," *IEEE Trans. on ASLP*, Vol. 16, No. 2, pp. 255–266, 2008.
 - [10] R. Macrae and S. Dixon: "Guitar tab mining, analysis and ranking," *ISMIR*, pp. 453–458, 2011.
 - [11] R. Macrae and S. Dixon: "A guitar tablature score follower," *ICME*, pp. 725–726, 2010.
 - [12] P. D. O'Grady and S. T. Rickard: "Automatic hexaphonic guitar transcription using non-negative constraints," *ISSC*, pp. 1–6, 2009.
 - [13] M. Paleari, B. Huet, A. Schutz, and D. Slock: "A multimodal approach to music transcription," *IEEE ICIP*, pp. 93–96, 2008.
 - [14] H. Penttinen, J. Siiskonen, and V. Valimaki: "Acoustic guitar plucking point estimation in real time," *IEEE Trans. on ASLP*, Vol. 3, pp. 209–212, 2005.
 - [15] K. Yoshii and M. Goto: "A nonparametric Bayesian multipitch analyzer based on infinite latent harmonic allocation," *IEEE Trans. on ASLP*, Vol. 20, No. 3, pp. 717–730, 2012.
 - [16] Guitar Warrior – Printable guitar chords – guitar chords chart. 2011[Online]. Available: <http://www.guitar-warrior.com/Printable-guitar-chords-guitar-chords-chart.html>, website: (date last viewed 5 Apr. 2012).