

演奏記号を考慮したバイオリン運指の推定

長田 若奈^{1,a)} 酒向 慎司^{1,b)} 北村 正^{1,c)}

概要：我々は条件付き確率場を用いた習熟度に対応したバイオリン運指推定手法を提案してきた。しかし、推定運指は運指の自然さや演奏表現の適切さが不十分である問題があった。本論文では、従来よりも多くの楽譜情報を用いて素性の設計を行う。運指に関連する素性には様々なものが考えられるが、素性重みの学習に L1 正則化を用いることで、運指推定に寄与する重要な素性を判別する。実験では、基本的な素性だけをを用いた従来法に対して教本運指との一致率が向上することを確認し、また、推定された運指の自然さに関してバイオリン経験者による主観評価を実施したところ提案法による有効性が確認できた。

1. はじめに

バイオリンは一つの楽譜に多数の運指が対応するため、演奏時に運指を決定する必要がある。初級者にとっては適切な運指を判断することが難しく、中級者以上であっても試行錯誤が必要となることがある。そのため、楽譜から適切な運指を自動推定する技術は、幅広い演奏者にとって有益である。また、運指は正確に演奏するための物理的な動きだけでなく、演奏の表現にも関わる重要な要素であるため、演奏分析や自動演奏生成にも寄与できると考えられる。

一般的に、演奏者によって適切な運指が異なることがある。これは、演奏者の身体的特徴や、演奏の癖、習熟度、楽曲解釈など様々な要因が影響している。自動運指推定においても、この点に注意する必要がある。本研究では特に習熟度による違いに着目する。同じ演奏者であっても、習熟度の向上により適切な運指は変化する。そのため、演奏技量の向上に合わせて運指を提示したり、利用者の要望に合わせて運指を作り分けるような技術によって、より効率的な練習支援システムが期待できる。このようなことから、本研究では任意の習熟度に対応した運指を推定することを目的とする。

今まで我々は任意の習熟度に対応した運指推定法を提案してきた [1]。これは、楽譜に対応する適切な運指との関係を確認率モデルによって記述するもので、演奏の容易さと演奏表現の優先度を習熟度によって変化させることで習熟度に対応した運指推定を可能とした。これにより、習熟度への対応は確認されたものの、演奏表現の適切さや運指の自然さでは教本運指に対して劣る結果となった。従来法では、楽譜の情報として音高と音符長と休符長のみを用いて

いるが、本来、演奏表現とは様々な楽譜情報から判断されるものである。また、バイオリンでは音階に基づく指の配置の方が適切とされているが、それも考慮されていないなどの問題が残されている。

一方で、一つの楽譜に多数の運指が対応する楽器はバイオリンだけでなく、今まで、様々な楽器で運指推定が行われてきた。しかし、多くの研究では音高情報のみから運指を推定している。音高以外の情報にも着目した運指推定の研究は、休符長から運指の連続性が失われる部分を考えるもの [2]、音符長と休符長からフレーズ分割を行うもの [3]、事前に付与されたフレーズ境界のアノテーションを用いるもの [4]、発音の時間間隔を運指の遷移の適切さに取り入れたもの [5] などがある。しかし、これらの研究では演奏表現にまで言及しているものは無く、演奏表現を考慮した運指推定にどのような楽譜情報を用いるべきかは定まった知見がない状況である。

本研究では、従来法における課題を解決するため、我々が提案した条件付き確率場 (CRF) に基づく習熟度に対応した運指推定 [6] を拡張し、より多くの楽譜情報を用いた素性を追加する。ここでの楽譜情報とは、基本的には楽譜に記載されている記号的な情報を対象とし、何らかの解釈を要する高次元情報は扱わないものとする。その場合、楽譜に記載されている多様な情報から、運指推定に寄与するものを区別できることが望ましい。そこで、素性重みの学習に L1 正則化を取り入れることにより、不要な素性の重みを 0 にする。これによって、重みの学習時に必要な素性と不要な素性を判別する。実験では、教本運指との比較やバイオリン経験者による主観評価実験により、演奏表現の適切さと運指の自然さが向上するか検証する。

2. CRF に基づく習熟度に対応したバイオリン運指推定

ここでは、これまでに提案した従来研究 [6] について簡単に述べる。

¹ 名古屋工業大学大学院
Graduate School of Engineering, Nagoya Institute of Technology, Gokiso, Showa, Nagoya, Aichi, 466-8555 Japan
a) nagata@mmsp.nitech.ac.jp
b) sako@mmsp.nitech.ac.jp
c) kitamura@mmsp.nitech.ac.jp

2.1 CRFに基づく運指のモデル化

運指推定を楽譜 o から音符毎に遷移する運指系列 f を求める問題と考える。ただし、ある時刻に発音されている音は1つとし、重音や多声部を含む楽譜は考えない。本研究では、運指にマルコフ性を仮定し、楽譜と運指の対応関係をCRFでモデル化することで系列ラベリング問題として解く。

CRFの観測変数は楽譜 o とし、潜在変数は運指をより詳細化した手の状態 s とする。手の状態 s は運指に対応する弦と指に加え、手の位置と指の間隔によって表し、これらの組み合わせとして式(1)と定義する。

$$s = \{x_S, x_F, x_P, x_I\} \quad (1)$$

ここで、 x_S は弦、 x_F は指番号、 x_P は手の位置(人差し指のうちのフレット番号)、 x_I は指の間隔(各指の間隔が半音か全音か)である。

楽譜 o に対する手の状態系列 s の適切さは素性 ϕ とその重み w によって式(2)のように表される。

$$P(s|o) = \frac{1}{z(o)} \exp \left\{ \sum_{n,i} w_i \phi_i(s_n, s_{n-1}, o, n) \right\} \quad (2)$$

ここで、 $z(o)$ は正規化項であり、式(3)である。

$$z(o) = \sum_s \exp \left\{ \sum_{n,i} w_i \phi_i(s_n, s_{n-1}, o, n) \right\} \quad (3)$$

与えられた楽譜 o に対する最尤状態系列 \hat{s} が最適運指であり、式(4)のように表される。

$$\hat{s} = \arg \max_s P(s|o) \quad (4)$$

2.2 習熟度への対応

本研究では、習熟度による運指の違いを、演奏の容易さと演奏表現のうち、どれほど演奏表現を優先するのかの違いであると考えられる。例えば、習熟度が低い演奏者は最も容易な運指でしか演奏できず、演奏表現が高い難しい運指を避け演奏の容易な運指を選択する。一方、習熟度の高い演奏者は難しい運指も演奏できるため、演奏表現に合った運指を選択すると考えられる。

また、演奏表現の優先度の違いは音符長にも依存すると考えられる。同じ習熟度であっても、音符長が短い場合は演奏が困難になりやすく、演奏の容易さの方が優先されやすい。一方、音符長が長い場合は比較的演奏が容易になりやすく、演奏表現が優先されやすい。また、音符長が長い音の方が聴者にとって耳に残りやすく重要であるため、演奏表現が優先されるべきである。

本研究では、演奏の容易さと演奏表現のうち、どれほど演奏表現を優先するのかを表現度 e と呼び、式(5)と定める。

$$e_n = L \log(1 + l_n) \quad (5)$$

ここで、音符長 l についての変化に対数を用いているのは、音符長が長い場合よりも短い場合の方が音符長による表現度の変化が大きいと考えられるからである。

このように習熟度 L と音符長 l の両方から表現度 e を定義することによって、単一の習熟度の運指から異なる習熟度の運指を考えることができる。習熟度 L が大きい演奏者の運指であっても、音符長 l が短い場合は e が小さくなり、

習熟度 L の低い演奏者の運指の傾向に近づく。よって、あらゆる習熟度の運指を用意することなく、任意の習熟度の運指を推定することを可能とする。

2.3 素性

状態に関する素性

指、手の位置、指の間隔について、それぞれ出現確率の対数を素性とする。ただし、出現確率はゼロ頻度問題を解決するため、ラプラススムージングを行う。また、指番号0については開放音(何も押さえないで発せられる音)のみでしか用いないため、開放音の出現確率の逆数で重み付けをし、出現確率を調整する。

状態遷移に関する素性

一般的に、長い音の後や休符では、フレーズ境界である可能性が高く、また手の移動を行う時間的余裕があるため、状態間の遷移が起こりやすい。このような、楽譜による状態のvarietyやすさを本研究では可変度 c と呼び、式(6)のように定める。

$$c_n = l_n + Rr_n \quad (6)$$

弦、手の位置、指の間隔の遷移について、移動量が分散が可変度に依存するようなラプラス分布や指数分布であると仮定し、それらの分布の対数を素性とする。

状態と音高に関する素性

音高に対応する状態の場合は0、そうでない場合は $-\infty$ とする。

状態と表現度に関する素性

弦と指について、それぞれ表現度の分布を対数正規分布と仮定し、その分布の対数を素性とする。

3. 新たな素性設計

3.1 ポジションの考慮

3.1.1 バイオリンにおける運指のポジション表現

一般に、運指列は音階に従っている事が多い。また、演奏の際には弦に対する水平方向の手の位置を考慮することが重要である。このような事から、それぞれの手の位置について音高に対する弦と指の組み合わせのテンプレートが考えられ、そのテンプレートに基づいた手の位置はポジションと呼ばれている。

ポジションは、指番号1(人差し指)の位置の音と開放音の音程が2度の時は第1ポジション、3度のときは第2ポジション、4度のときは第3ポジション、...というように定められている。ギターなどのフレット表現は半音単位であるのに対し、バイオリンのポジション表現は度単位であることに注意が必要である。各ポジションのテンプレートは、指番号1, 2の音の音程が2度、指番号1, 3の音の音程が3度、指番号1, 4の音の音程が4度、となるように定



図1 第3ポジションにおける音高に対する弦と指

Fig. 1 String and finger according to pitch in 3rd position

められている．指番号 0 はテンプレートに定義されていない．例として，第 3 ポジションのテンプレートを図 1 に示す．変化記号*1 がある場合も変化記号がない場合と同じである．音高（音名とオクターブ）とポジションから弦と指を考慮することができ，ポジションによって間接的に運指を表すことができる．また，音高（音名とオクターブ）と弦と指からポジションを考慮することができる．

一般に，バイオリン演奏者はポジションの概念で手の位置を考え，運指を判断することが多い．このため，ポジションに関する素性を組み込むことで，より自然な運指が推定できると考えられる．ただし，ポジションは音階に従って指が配置されているという前提がある．多くの場合は音階に従って指を配置するが，必要に応じて音階に従う運指から外れることがある．このため，常に手の位置を考慮することができるよう，従来の手の位置の素性も必要である．

3.1.2 ポジションに関する素性の追加

楽譜と運指からポジションを考え，ポジション自体とポジション遷移の適切さに関する素性を加える．指番号 0 はポジションで定義されていないが，前後の音が第 1 ポジションの時に使われることが多いため，第 1 ポジションとする．

ポジション自体に関する素性

ポジションの出現確率の対数を素性とする．状態自体についての素性と同様に，出現確率はラプラススムージングを行う．

ポジション遷移に関する素性

状態遷移と同様に，移動量が分散が可変度に依存するようなラプラス分布であると仮定し，それらの分布の対数を素性とする．

3.2 演奏記号の考慮

3.2.1 運指と演奏記号の関係

演奏記号とは，五線や拍子，音符，休符などを除いた，記譜法上必要不可欠でない記号のことであり，強弱や速度，アーティキュレーション（音の形）などを示す記号である．文字によって示されるものとマークやシンボルによって示されるものがあるが，いずれも記号と呼ぶ．演奏記号は楽曲の演奏表現にとって重要であり，運指にも影響を与えると考えられる．よって，演奏記号と運指の対応を素性に組み込む事を考える．

ただし，演奏記号は種類が多く，また楽曲に依存して出現頻度に偏りがあるため，全ての演奏記号を扱うことは困難である．特に，言葉によって演奏表現を指示しているものは，個別に扱うには語数が多く，言葉の意味の解釈が必要となる．よって本研究では原則的にマークやシンボルで表される演奏記号のみを扱う．ただし，速度や音量を指示する言葉は限られているため，それらはテンポや音量の数値に置き換えて用いる．また，Sul 記号は作曲者による弦の指定であり，運指に重要と考えられるため，Sul 記号も扱う．

3.2.2 表現度と演奏記号の関係

演奏記号は演奏表現を考慮する場合に必要となるが，逆に，運指の容易さのみを考えるには不要である．このよう

表 1 強弱の定義

Table 1 Definition of dynamics

| | | | |
|---------------|---|-------------------------|--------------|
| <i>ppp</i> 以下 | 1 | <i>ff</i> | 7 |
| <i>pp</i> | 2 | <i>fff</i> 以上 | 8 |
| <i>p</i> | 3 | <i>piu f</i> | 前の強弱 +1 |
| <i>mp</i> | 4 | <i>piu p</i> | 前の強弱 -1 |
| <i>mf</i> | 5 | <i>sf sfz fz rfz rf</i> | その音のみ前の強弱 +2 |
| <i>f</i> | 6 | <i>fff fffz fffz</i> | その音のみ前の強弱 +3 |

なことから，表現度 e から演奏記号の考慮度を考え，それを素性に重み付けする．演奏記号の考慮度 C は式 (7) のように定める．

$$C_i(o, n) = \begin{cases} \frac{e_n}{\bar{e}} & \text{状態についての素性} \\ \frac{(e_{n-1} + e_n)/2}{\bar{e}} & \text{状態遷移についての素性} \end{cases} \quad (7)$$

ここで， \bar{e} は学習データの e の平均である．この C を素性に重み付けし，最終的な素性関数を式 (8) とする．

$$\phi_i(s_n, s_{n-1}, o, n) = C_i(o, n) \psi_i(s_n, s_{n-1}, o, n) \quad (8)$$

以降，演奏記号の考慮度 C による重み付けをする前の素性 ψ を原素性と呼ぶこととする．

3.2.3 状態と演奏記号に関する素性の追加

ある 1 つの状態と関係のある楽譜情報は，その音の強弱，テンポ，音符長，その音に付随する演奏記号，現在より前に出現した Sul 記号であると考えられ，全て独立であるとする．また，演奏表現に関係する状態要素は指と弦のみと考える．

トリルによる指の制約の素性

トリルでは指番号 4（小指）になることはない．このため，トリルの音で指番号が 4 になる場合は $-\infty$ ，それ以外は 0 とする．

状態と強弱・テンポに関する素性

強弱については表 1 のように数値化する．クレッシェンドなどの徐々に変化するものについては，次の強弱記号が出てくるまで線形変化するとする．次の強弱記号がない，もしくは強弱変化が記号の示すものと反対の場合は，変化開始時の強弱の -1 もしくは $+1$ までの変化とする．

テンポについては，テンポ (BPM) が直接指定されているもの以外に，速度標語で示されているものがある．速度標語については，標準的な BPM に変換して考える．相対的なものは，*meno mosso* は前のテンポの 90%，*piu mosso* は前のテンポの 110% とする．徐々に変化するものについては，次のテンポ記号が出てくるまで線形変化するとする．次のテンポ弱記号がない，もしくはテンポ変化が記号の示すものと反対の場合は，変化開始時のテンポの 90% もしくは 110% までの変化とする．*molto*（とても）がついているものは変化量を 1.5 倍，*poco*（少し）がついているものは変化量を 0.5 倍とする．

各指，弦に対し，強弱，テンポの分布は正規分布とし，その分布の対数を原素性とする．

状態と音符長に関する素性

音符長は演奏記号には含まれないが，演奏表現に寄与すると考えられるため，演奏記号と同様に考えて扱う．各指，弦に対し，音符長の分布が対数正規分布であると仮定し，

*1 シャープやフラットなど

その分布の対数を原素性とする。

状態と音符に付随する演奏記号に関する素性

演奏記号の有無を $m = \{0, 1\}$ で表し, $m = 0$, $m = 1$ それぞれにおける状態要素の出現確率 $P(x|m)$ を考える。ただし, これには状態要素自体の出やすさも含まれているため, 演奏記号による状態の出現確率の変化のみを考えるために状態要素の出現確率で割る。これを用いて, 原素性を式 (10) とする。

$$\psi(s_n, \mathbf{o}, n) = \log \frac{P(x_n|m_n)}{P(x_n)} \quad (9)$$

ここで, $P(x_n|m_n)$, $P(x_n)$ は状態自体についての素性と同様の理由から, ラプラススムージング, 指番号 0 の出現確率の調整を行う。

状態と Sul 記号に関する素性

Sul 記号とは, 作曲者によって演奏弦を指定する記号であり, 例えば “Sul G” と記載されていれば G 線で演奏することを示している。Sul 記号は適用範囲の最初のみ書かれ, 適用範囲の終端は演奏者が作曲者の意図を読み取って判断する。このような判断は高度な楽譜解釈が必要であるため, 本研究では, 簡易的に, Sul 記号の記載位置から遠くなるほど適用されている可能性が低くなることを考え, Sul 記号の適用度を $S_{\text{Sul}}(\mathbf{o}, n)$ とする。Sul 記号の記載位置を 1 とし, 16 拍で 0 となるように線形減少するような関数である。原素性は, $S_{\text{Sul}}(\mathbf{o}, n)$ の対数を考えるが, $-\infty$ とならないよう, 最低値を k_{sul} として次のように定める。

$$\psi(s_n, \mathbf{o}, n) = \begin{cases} \max(\log S_{\text{Sul}}(\mathbf{o}, n), k_{\text{sul}}) & x_{S,n} = S_{\text{Sul}}(\mathbf{o}, n) \\ k_{\text{sul}} & x_{S,n} \neq S_{\text{Sul}}(\mathbf{o}, n) \end{cases} \quad (10)$$

ここで $S_{\text{Sul}}(\mathbf{o}, n)$ は Sul 記号によって指定された弦である。

3.2.4 状態遷移と演奏記号に関する素性の追加

状態遷移に関係のある楽譜情報は, 遷移前後の強弱, テンポの変化, スラー, 装飾音符かどうか, フレーズ境界であると考えられ, 全て独立であるとする。また, 演奏表現に関係する状態要素は弦と手の位置と指の間隔のみと考える。

状態遷移と強弱・テンポ変化に関する素性

強弱やテンポの変化がある部分は曲の雰囲気が変わり, 運指の傾向も変わりやすいと考えられる。よって強弱やテンポの変化の有無と状態遷移を対応付ける。原素性は, 状態と音符に付随する演奏記号についての素性と同様に考え, 式 (11) とする。

$$\psi(s_n, s_{n-1}, \mathbf{o}, n) = \log \frac{P(t_x(x_n, x_{n-1})|t_m(o_n, o_{n-1}))}{P(t_x(x_n, x_{n-1}))} \quad (11)$$

ここで $t_x(x_n, x_{n-1})$ は状態遷移についての関数であり, 状態の変化があったかどうかを 0, 1 で表す。また, $t_m(o_n, o_{n-1})$ は強弱やテンポの変化があったかどうかを表し, 強弱は変化量が k_{dy} 以下の場合, テンポは変化率が $1 - k_{\text{tempo}}$ 以上 $1 + k_{\text{tempo}}$ 未満の場合は変化がないとみなす。

状態遷移とスラーに関する素性

遷移前後の音がスラーで繋がっているかどうかは状態の

遷移に関わると考えられる。原素性は, 状態遷移と強弱・テンポ変化についての素性と同様に考え, 式 (12) とする。

$$\psi(s_n, s_{n-1}, \mathbf{o}, n) = \log \frac{P(t_x(x_n, x_{n-1})|m(\mathbf{o}, n))}{P(t_x(x_n, x_{n-1}))} \quad (12)$$

ここで, $m(\mathbf{o}, n)$ は o_{n-1} と o_n がスラーで繋がっているかどうかを表す。

状態遷移と装飾音符に関する素性

装飾音符は演奏記号には含まれないが, 音符が装飾音符かそうでないかが状態遷移に影響を及ぼすと考えられる。よって遷移前の音符が装飾音符かどうかを考える。 $m(\mathbf{o}, n)$ を o_{n-1} が装飾音符かどうかを表すものとする, 原素性は式 (12) と同じである。

状態遷移とフレーズ境界に関する素性

先行研究 [3], [4] で運指推定にフレーズ境界を用いているように, フレーズの境界と運指には関係があると考えられる。フレーズ分割には高度な楽譜解釈が必要であるが, 休止記号やプレス記号, フェルマータはフレーズ終端についていると考えられ, このような記号がある部分については高度な楽譜解釈なしにフレーズ境界を判断できると考えられるため, これを素性に取り入れる。 $m(\mathbf{o}, n)$ を o_{n-1} にこれらの記号があるかどうかを表すものとする, 原素性は式 (12) と同じである。

4. L1 正則化 SGD による素性重みの学習

本提案では, 新たに運指に関係がある可能性のある素性を加えた。しかし, 演奏表現は右手の弓の動きによるところが大きく, 運指との関係の有無やその度合いは定かではない。また, モデルの過度な複雑化を避ける必要もあるため, 必要な素性と不要な素性を自動的に判別することを考える。

素性重みは各素性の重要度を表すため, 不要な素性は素性重みが 0 になる事が望ましい。これは, L1 正則化によって実現できると考えられる。L1 正則化では常に 0 になる方向にペナルティが働くため, 不要な素性重みを積極的に 0 にすることができる。従来法 [6] では, 正則化なしの確率的勾配降下法 (SGD) を用いて素性重みの学習を行っていた。本研究では, SGD に L1 正則化を取り入れる。L1 正則化を取り入れた重み更新式は式 (13) となる。

$$w_i^{k+1} = w_i^k + \eta_k \left\{ \phi_i(\mathbf{o}_k, \mathbf{s}_k) - \sum_{\mathbf{s}} \phi_i(\mathbf{o}_k, \mathbf{s}) P(\mathbf{s}|\mathbf{o}_k) - \frac{\partial}{\partial w_i} \frac{C}{N_d} |w_i| \right\} \quad (13)$$

ここで, k は更新回数, N_d は学習データ数, C は正則化の強さを表す定数である。本研究では, Tsuruoka ら [7] の手法で式 (13) を解く。

なお, 状態と音高についての素性とトリルによる指の制約の素性は 0 と $-\infty$ のみを取るため, 重みの学習は不要である。これらの素性は必ず必要であるため, 0 より大きい任意の値とする。

5. 評価実験

5.1 パラメータの学習

用いたデータを表 2 に示す。用いた教本は 3 種

表 2 実験に用いたデータ

Table 2 Details of experimental data collection

| | | 習熟度 | 曲数 | 音符数 | 運指記載数 |
|-----|------|-----|----|--------|-------|
| 学習 | 手動補完 | 中級 | 14 | 4,550 | 641 |
| | 自動補完 | 中級 | 71 | 25,800 | 4,347 |
| テスト | | 初級 | 13 | 1,906 | 189 |
| | | 中級 | 13 | 4,890 | 1,295 |

類 [8], [9], [10] であり, いずれのデータも重複しない. 手動補完データは, バイオリン経験者が教本運指 (部分的に運指が記載されているもの) を元に全音符についての運指を付与したものである. 自動補間データは, 先行研究 [6] の方法で手動補完データを学習データとして教本運指を補完したものである. ただし, 手動補完データは運指補完モデルの学習にとって十分な量であるとし, 先行研究 [6] における補完の繰り返しによる再学習は行っていない. また, 補完用パラメータの学習時の SGD 終了条件は更新 100 回とした. 学習データは 50 音ずつに区切ったものであり, 学習には表 2 で示した手動補完データ, 自動補間データから, ランダムに 200 区間 (10,000 音) を用いた.

SGD における学習率は $\eta_t = 0.01 \times 0.5^{t/200}$, 正則化の強さは $C = 20$ とし, 終了条件は更新 200 回とした. 重みの初期値は学習が不要な 2 つの素性以外は 0 とした. 素性中のパラメータは経験的に $k_{sul} = \log 10^{-3}$, $k_{dy} = 1.0$, $k_{tempo} = 0.1$ とした. また, 可変度についてのパラメータは $R = 4.0$, 学習時の習熟度 L は 1.0 とした.

学習に用いた素性と推定された素性重みを表 3 に示す. 重みの学習が必要な素性 42 個のうち, 24 個が重み 0 になった.

5.2 教本運指との一致率による評価

5.1 節で学習したモデルを用いて, 表 2 のテストデータに対して運指推定を行い, 教本運指から人手で補完したものと的一致率を調べた. 推定時の習熟度 L は 0.1 から 3.0 まで 0.1 刻みとした.

結果を図 2 に示す. 全素性の場合も従来の素性の場合も初級テストデータと中級テストデータでピークの位置が異なっており, 本提案の手法でも段階的な習熟度に対応していると言える. 初級テストデータでは, 全素性では $L = 0.3$ で最大値 94.5% であり, 従来の素性のみでは $L = 0.4$ で最大値 91.4% である. 中級テストデータでは, 全素性では $L = 1.1$ で最大値 76.1% であり, 従来の素性のみでは $L = 0.9$ で最大値 75.0% である. それぞれのテストデータにおいて, 全素性の場合と従来の素性の場合の最大値を McNemar 検定をしたところ, $P < 0.05$ で有意差があった. このようなことから, ポジションや演奏記号を考慮することにより, 運指推定精度が向上し, また, 従来研究のような習熟度への対応も保たれていると言える.

初級テストデータでは高い一致率が得られており, 学習データが中級のみであっても初級向き運指が推定できると言える. 一方, 中級テストデータでは一致率は 76% 程度であるが, 最適な運指には個人差があるため, 一致していない部分が一概に悪い運指とは言えない場合もある. このようなことから, 中級テストデータについては主観評価実験を行う.

表 3 学習に用いた素性と推定された素性重み

Table 3 List of features and estimated feature weights

| 素性 | 素性重み |
|--------------------|-------|
| 状態自体 (手の位置) | 0.574 |
| 状態自体 (指の間隔) | 0.321 |
| 状態遷移 (弦) | 0.900 |
| 状態遷移 (手の位置) | 1.067 |
| 状態遷移 (指の間隔) | 0.744 |
| 状態と表現度 (指) | 0.341 |
| 状態と表現度 (弦) | 0.155 |
| ポジション自体 | 0.096 |
| ポジション遷移 | 0 |
| 状態と音量 (指) | 0 |
| 状態と音量 (弦) | 0.009 |
| 状態とテンポ (指) | 0.097 |
| 状態とテンポ (弦) | 0.066 |
| 状態と音符長 (指) | 0.491 |
| 状態と音符長 (弦) | 0.551 |
| 状態とアクセント (指) | 0 |
| 状態とアクセント (弦) | 0 |
| 状態とスタッカート (指) | 0 |
| 状態とスタッカート (弦) | 0 |
| 状態とテヌート (指) | 0 |
| 状態とテヌート (弦) | 0 |
| 状態とフェルマータ (指) | 0 |
| 状態とフェルマータ (弦) | 0 |
| 状態とトリル (指) | 0.146 |
| 状態とトリル (弦) | 0 |
| 状態と Sul 記号 (弦) | 0 |
| 状態遷移と音量変化 (弦) | 0 |
| 状態遷移と音量変化 (手の位置) | 0.120 |
| 状態遷移と音量変化 (指の間隔) | 0 |
| 状態遷移とテンポ変化 (弦) | 0 |
| 状態遷移とテンポ変化 (手の位置) | 0.004 |
| 状態遷移とテンポ変化 (指の間隔) | 0 |
| 状態遷移とスラー (弦) | 0.378 |
| 状態遷移とスラー (手の位置) | 0 |
| 状態遷移とスラー (指の間隔) | 0 |
| 状態遷移と装飾音符 (弦) | 0 |
| 状態遷移と装飾音符 (手の位置) | 0 |
| 状態遷移と装飾音符 (指の間隔) | 0 |
| 状態遷移とフレーズ境界 (弦) | 0 |
| 状態遷移とフレーズ境界 (手の位置) | 0 |
| 状態遷移とフレーズ境界 (指の間隔) | 0 |

5.3 バイオリン経験者による主観評価

バイオリン経験者 6 名 (経験歴 8~18 年) による主観評価を行った. 被験者には, 1 つの楽譜につき 3 つの運指を示し, 各運指について 3 つの項目について評価するよう求めた. 提示した運指は 5.2 節における全素性と従来素性の場合の $L = 1.0$ での推定運指と教本運指であり, 運指の順序はランダムである. 曲は, 中級テストデータの各曲において運指に最も違いのある 4 小節を含むように 478 小節を抜粋した. 評価項目は次の 3 つであり, 回答は各 5 段階である.

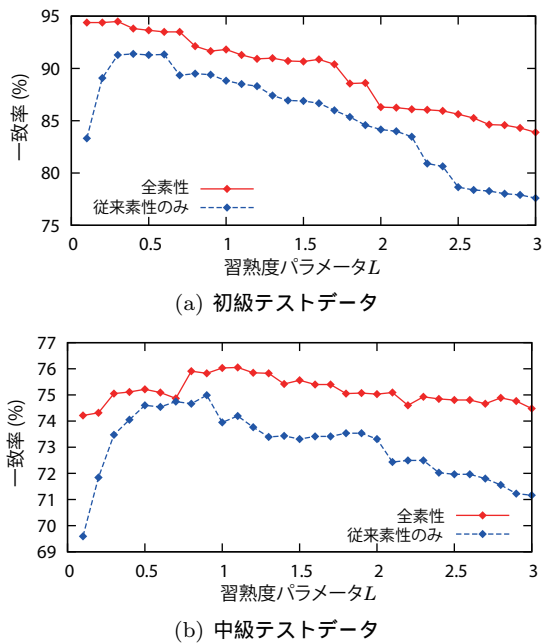


図 2 教本運指との一致率
Fig. 2 Rate of concordance with textbook fingering

容易さ 演奏表現などを考慮しない場合に演奏が容易であるか (1 難しい - 3 どちらとも言えない - 5 容易)
表現の適切さ 演奏表現を考慮した場合に運指が適切であるか (1 不適切 - 3 どちらとも言えない - 5 適切)
自然さ 運指が自然かどうか (1 不自然 - 3 どちらとも言えない - 5 自然)

結果を図 3 に示す。従来素性のみの場合と比較し、全素性の場合では容易さと自然さで改善が見られた。一方、表現の適切さでは有意な差がなかった。これは、演奏表現にはより高度な楽譜解釈が必要であり、単純な演奏記号の考慮では表現の適切さを改善するには不十分だったと考えられる。しかし、表 3 でポジションに対する素性重みが比較的小さいことを考えると、単純な演奏記号の考慮は自然さの向上に貢献したと考えられるため、演奏記号を素性に用いる有効性が示されたと言える。

一方、教本補完運指と比較すると、いずれの項目でも同等以上の結果となった。よって、一致率では 76% であっても、本研究により推定された運指は、潜在的には一定の実用性があることが示唆された。教本を上回る評価が得られた結果については、さらに評価結果を精査する必要があるが、考えられる理由としては、個人によって最適な運指が異なることが挙げられる。教本にはある程度良い運指が記載されていると考えられるが、推定運指は学習データに 3 種類の教本を用いているため、より一般的な運指となり、教本以上の評価が得られたと考えられる。

6. おわりに

本研究では、従来の CRF によるバイオリン運指モデルを拡張し、ポジションおよび演奏記号の素性を加えた。また、素性重みの学習に L1 正則化を取り入れることによって必要な素性の選別を行った。実験では、約半数の素性重みが 0 になる事を確認し、教本運指との一致率の向上も確認された。主観評価実験では、演奏の容易さと自然さの改

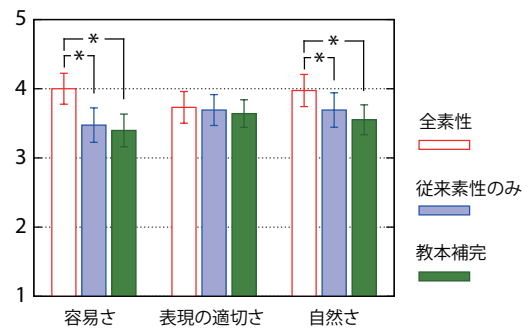


図 3 主観評価結果 (エラーバーは 95% 信頼区間, * は有意水準 5% の t 検定において有意差ありの組)

Fig. 3 Results of the subjectivity evaluation (The error bar represents the 95% confidence interval and the asterisk (*) represents a pair that has significance difference between the two evaluations in t-test at 5% significance level).

善が示された。

本研究では表層的な演奏記号を素性の設計に用いたが、より楽曲に合った適切な運指を得るには高度な楽譜解釈に基づく演奏表現の考慮も必要である。

謝辞 本研究は JSPS 科研費 (26730182) および財団法人電気通信普及財団の助成を受けた。

参考文献

- [1] W. Nagata, S. Sako, and T. Kitamura: Violin Fingering Estimation According to Skill Level based on Hidden Markov Model, in *Proc. ICMC and SMC*, 1233-1238 (2014).
- [2] 春原雅彦, 松原正樹, 前田俊太郎: 左手と和音に対応したピアノ運指の自動推定システム, 情報処理学会全国大会講演論文集 (2), pp. 147-148 (2010).
- [3] M. Miura, I. Hirota, N. Hama, and M. Yanagida: Constructing a system for finger-position determination and tablature generation for playing melodies on guitars, *System and Computers in Japan*, vol. 35, no. 6, pp. 10-19 (2004).
- [4] D. Radicioni, L. Anselma, and V. Lombardo: A segmentation-based prototype to compute string instruments fingering, in *Proc. the Conference on Interdisciplinary Musicology* (2004).
- [5] G. Hori, H. Kameoka, and S. Sagayama: Input-Output HMM Applied to Automatic Arrangement for Guitars, *Journal of information processing*, vol. 21, no. 2, pp. 264-271 (2013).
- [6] 長田若奈, 酒向慎司, 北村正: 条件付き確率場に基づくバイオリン運指推定のための教本からのパラメータ学習, 日本音響学会 2014 年秋季研究発表会, pp. 909-912 (2014).
- [7] Y. Tsuruoka, J. Tsujii, and S. Ananiadou: Stochastic gradient descent training for L1-regularized log-linear models with cumulative penalty, in *Proc. ACL and AFNLP*, pp. 477-485 (2009).
- [8] 鈴木鎮一: 鈴木バイオリン指導曲集 2, 4, 5, 6, 全音楽譜出版社 (1971-2009).
- [9] 篠崎弘嗣: 篠崎バイオリン教本 2, 3, 4, 全音楽譜出版社 (1955-1987).
- [10] 兎束龍夫, 篠崎弘嗣, 鷲見三郎: 新しいバイオリン教本 3, 4, 音楽之友社 (2012-2013).