E-029

習熟度に応じたバイオリン運指推定のための確率モデルとパラメータ学習

Statistical Modeling and Parameter Learning for Violin Fingering Estimation Corresponding to Skill Level

長田 若奈 † Wakana Nagata
酒向 慎司 † Shinji Sako
北村 正 † Tadashi Kitamura

1. はじめに

バイオリンは一つの楽譜に多数の運指が考えられる楽器である。教本以外の楽譜には運指は記述されていないため演奏者自身で運指を決定する必要があるが，運指の決定には経験や試行錯誤が必要である。このため，運指を自動推定する必要があるが，適切な運指は演奏者によって異なる。

我々は，習熟度による運指の違いは演奏表現の優先度合いの違いであると考え，習熟度に応じた確率モデルに基づいた運指推定法を提案した [1]。しかし，運指の適切さであるモデルパラメータを経験的に設定する必要があり，特に演奏表現を優先する場合ではモデルが複雑化するため適切に設定することは困難であった。

そこで本研究では，運指推定モデルにおける運指の適切さを，教本の運指データを用いて学習する手法を提案する。提案法では音符長や休符長に依存して変動する結合度，表現度について出現頻度分布を仮定することで，これらに依存する確率を定める。また，過学習を防ぐためにスムージングを行う。

2. 習熟度に応じたバイオリン運指の推定

2.1 バイオリン運指のモデル化

ある音符が演奏される際に考えられる左手の状態 s を，SP (String Position)，FN (Finger Number)，HP (Hand Position)，FI (Finger Interval) によって表わすこととする。各要素はそれぞれ 4, 5, 24, 8 個の値を取るとすると，状態数は 3840 である。

本研究ではある音符列に対応する運指を，このような状態の遷移として考え，状態の適切さは１つ前の状態のみに依存するとする。ここで，s_i から s_j への遷移の適切さ (遷移確率 $a_{i,j}$) と s_i の押弦の適切さ (押弦確率 b_i) により，運指の適切さを考えることが可能である。状態が遷移する際の状態の変化の容易さは音符に依存し，どれほど表現を優先するべきかは音符と習熟度に依存する。

2.2 可変度と表現度

運指の変化の容易さ (可変度 C) と演奏表現の優先度 (表現度 E) は音符長や休符長に依存すると考えられ，それぞれ $C_n = L_{n-1} + w_r R_{n-1,n}$，$E_n = \log_{base}(L_n + 1)$ と定める。ここで L_n は n 音目の音符長，$R_{n-1,n}$ は $n-1$ 音目と n 音目の間の休符長，w_r は休符長の重みである。対数の底 $base$ は習熟度を表し，値が大きくなるほど演奏表現が優先されにくく，習熟度の低い奏者向きの運指が推定される。

2.3 遷移確率，押弦確率

経験的に，要素間の依存関係は，遷移確率の HP と FI は FN に影響し，他は全て独立であると考えられる。また，可変度 C は遷移確率，表現度 E は押弦確率の SP と FN のみに影響を与えると考えられる。これらをふまえて，遷移確率，押弦確率をそれぞれ式 (1), (2) とする。

$$a_{i,j}(C) = P_{SP}(i,j,C)P_{FN|HP,FI}(i,j,C) P_{HP}(i,j,C)P_{FI}(i,j,C) \quad (1)$$

$$b_i(E) = P_{SP}(i,E)P_{FN}(i,E)P_{HP}(i)P_{FI}(i) \quad (2)$$

3. モデルパラメータの学習

前節で述べたパラメータを手動で設定するには，バイオリン演奏の知識や経験，試行錯誤が必要である。そこで，運指が併記された楽譜を用いて，楽譜から正解運指が推定されるようなパラメータを決定する手法を考える。

3.1 共通パラメータ

前節で述べたように，要素間の依存は限定的であると考えられる。しかし，遷移確率については，要素毎に考えても取りうる遷移の数は SP，FN，HP，FI それぞれの要素で 4^2, 5^2, 24^2, 8^2 個であり，多くのパラメータが存在する。そのため，経験的に各確率の依存関係を考え，いくつかの遷移は同じ確率とみなす。その集合を表 1 に示す。押弦確率については要素値ごとに考えるため，$X^{\alpha}_{x} = \{\alpha_x\}$ となる。

3.2 最尤推定

可変度 C，表現度 E に依存しない確率について，観測された状態または遷移の要素 α の重複集合 O^{α} を用いて，出現確率は式 (3) と表される。

$$P^{ML}_{\alpha}(i) = \frac{|\{\alpha_i\} \cap O^{\alpha}|}{|O^{\alpha}|} \quad (3)$$

ここで，$|X|$ は集合 X の要素の数である。

表 1: 同じ遷移確率とする遷移の集合

$X^{SP}_{x} = \{(SP_i, SP_j) \mid \lvert SP_i - SP_j \rvert = x\}$
$Y_{pq} = Y^{HP}_{p} \cap Y^{FI}_{q}$ $Y^{HP}_{p} = \{(i,j) \mid T(HP_i = HP_j) = p\}$ $Y^{FI}_{q} = \{(i,j) \mid T(FI_i = FI_j) = q\}$ $X^{FN}_{x\mid pq} = \begin{cases} \{(FN_i, FN_j)\} & (p,q) = (0,0) \\ \{(FN_i, FN_j) \mid T(FN_j = 0) = x\} & (p,q) = (0,1) \\ \{(FN_i, FN_j) \mid FN_j = x\} & p = 1 \end{cases}$
$X^{HP}_{x} = \{(HP_i, HP_j) \mid \lvert HP_i - HP_j \rvert = x\}$
$X^{FI}_{x} = \{(FI_i, FI_j) \mid T(FI_i = FI_j) = x\}$

$T(A)$: 命題 A が真なら 0, 偽なら 1 を返す関数

†名古屋工業大学大学院工学研究科，Graduate School of Engineering, Nagoya Institute of Technology

可変度 C に依存する確率について，集合 X_x^α の要素の出現頻度を $f(X_x^\alpha, C)$ とすると，集合内の要素の出現確率は等しいという仮定から，式 (4) となる。

$$P_\alpha^{ML}(i,j,C) = \frac{f(X_x^\alpha, C)}{|X_x^\alpha| \sum_n f(X_n^\alpha, C)} \quad (\alpha_i, \alpha_j) \in X_x^\alpha \tag{4}$$

遷移確率の FN については条件付き確率なので式 (5) となる。

$$P_{FN|HP,FI}^{ML}(i,j,C) = \frac{f(X_{x|pq}^{FN}, C)}{|X_{x|pq}^{FN}| \sum_n f(X_{n|pq}^{FN}, C)}$$
$$(FN_i, FN_j) \in X_{x|pq}^{FN}, (i,j) \in Y_{pq} \tag{5}$$

本研究では経験的に C の分布を対数正規分布と仮定し，出現頻度は式 (6) とする。

$$f(X_x^\alpha, C) = \frac{|X_x^\alpha \cap O^\alpha|}{\sqrt{2\pi\sigma_{\alpha_x}^2}C} \exp\left\{-\frac{(\log C - \mu_{\alpha_x})^2}{2\sigma_{\alpha_x}^2}\right\} \tag{6}$$

ここで，μ, σ^2 は $\log C_n$ $((\alpha_{n-1}, \alpha_n) \in X_x^\alpha)$ の平均と分散である。

表現度 E に依存する確率についても可変度 C と同様に考える。

3.3 スムージング

最尤推定では過学習が発生しやすく，特にゼロ頻度問題は避けるべきである。そこで，Dirichlet smoothing を行う。事前確率を P^P，その確かさを c とすると，スムージングをした確率は式 (7) となる。

$$P = \frac{|O|}{|O|+c}P^{ML} + \frac{c}{|O|+c}P^P \tag{7}$$

4. 評価実験

4.1 教本との一致率

推定運指と教本との一致率で提案法を評価する。SP と FN が一致するものを正解として正解率を算出する。実験に用いたデータを表 2 に示す。テストデータと学習データに用いた曲は重複しない。事前確率は表 2 以外の 1 曲 (音符 366 個, レベル 2) が正解率 100%となるよう経験的に付与した。また $w_r = 4$ とする。学習時の $base$ は 10^3 とし，推定時の $base$, c については $base = 10^n$ $(n = 1, 2, \ldots, 10)$, $c = 500 \times n$ $(n = 0, 1, \ldots, 6)$ でグリッドサーチした。

実験結果を表 3 に示す。事前確率を用いた場合を手動設定の場合としている。同じレベルのテストデータによるパラメータの学習と手動設定の正解率を McNemar 検定したところ，3 つのレベル全てにおいて有意差が認められた ($p < 0.05$)。レベルが上がるにつれ学習した場合の方が良い結果が得られた。これは，表現を優先した際の適切なパラメータの手動設定が困難である為だと考えられる。手動設定パラメータの信頼度 c を見ても，レベルが高くなると c が下がることが確認される。またレベルが上がるほど $base$ は下がるが，これは習熟度に対応していると言える。

表 2: 実験データ

学習データ	13 曲 (音符 5375 個, レベル 2)
テストデータ	レベル 1: 14 曲 (音符 2265 個)
	レベル 2: 8 曲 (音符 2372 個)
	レベル 3: 10 曲 (音符 2201 個)

表 3: 最大正解率とその時の条件

	最大正解率		
レベル	1	2	3
学習	94.26%	81.49%	64.15%
手動設定	96.87%	79.68%	59.52%

	base			*c*		
レベル	1	2	3	1	2	3
学習	10^{10}	10^3	10	3000	2500	1000
手動設定	10^{10}	10^6	10^2			

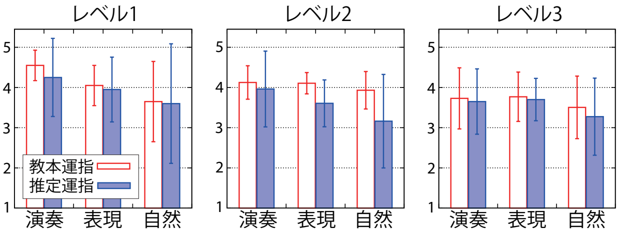

図 1: 主観評価の結果 (エラーバー：標準偏差)

4.2 バイオリン経験者による評価

適切な運指は教本運指のみではない為，教本との一致率のみでは運指の評価は充分ではない。そこで，バイオリン経験者 5 名（経験歴 3〜20 年）によって，同じ曲について推定運指と教本運指をそれぞれ評価した。被験者には教本運指か推定運指かは明かさず，楽譜と対応する運指を提示し，「演奏の容易さ」,「表現の容易さ」,「運指の自然さ」の 3 項目について 5 段階で評価させた。推定運指は 4.1 節の条件において各レベルで一致率が最大となった運指を用いた。評価には各曲の 10 小節目までを用いた。レベル 1 のうち 10 曲分は 10 小節間で一致しない運指が 1 つ以下のため除外した。

結果を図 1 に示す。結果から，推定運指においても各項目 3 点以上の結果を得ることが出来た。このことから，ある程度自然な運指が推定できていることが示唆される。レベル 2, 3 において特に運指の自然さで教本運指と推定運指の評価に差がある。表現の容易さと運指の自然さには相関がある事を考えると，より的確に演奏表現を捉える必要があると考えられ，そのためにはスラーや強弱等の他の運指に関わる楽譜情報も取り入れる必要があると考えられる。

5. むすび

本研究では，運指の適切さであるモデルパラメータを教本運指の出現確率として学習する手法を提案した。パラメータを手動設定した場合と学習した場合を教本との一致率で比較すると，習熟度が高い奏者向きの運指を推定する際に学習が有効であることを確認した。また，バイオリン経験者による主観評価実験では提案法である程度適切な運指を推定できることを確認した。しかし，習熟度が高い奏者向きの運指推定で評価が低い傾向があり，習熟度の向上による運指の変化をより的確に捉える事が今後の課題である。

参考文献

[1] 長田若奈, 酒向慎司, 北村正: “動的計画法に基づく音符長を考慮したバイオリン運指推定”, 情報処理学会第 75 回全国大会, pp.275–276, 2013.