

DP マッチングのコストに視線情報を反映させた 楽譜追跡手法の提案

寺崎 葉里^{1,a)} 竹川 佳成^{2,b)} 平田 圭二^{2,c)}

概要：本稿では DP マッチングのコストに視線情報を反映させた楽譜追跡手法を提案する。楽譜追跡は自動伴奏や自動譜めくりの要素技術として多く研究されてきた。従来の楽譜追跡の多くは演奏箇所の情報をもつ演奏の音響信号や MIDI などの打鍵情報をもとに演奏箇所を推定している。そのため、打鍵情報のみからでは現演奏箇所や演奏再開箇所の推定が難しい状況においては楽譜追跡に遅延や誤りが発生する。そこで、打鍵情報とは独立に演奏箇所情報をもつ視線を楽譜追跡結果に反映させることで解決を試みる。技術的な課題として、視線のノイズ低減や視線情報を反映させる手法の検討がある。そこで本研究では、HMM によって視線の動きを予測するモデル（視線モデル）を作成することで、視線のノイズ低減の効果を狙う。また、推定結果から視線の尤もらしさ（視線尤度）を算出し、視線尤度を DP マッチングのコストに反映させることで誤打鍵に左右されない楽譜追跡を実現する。提案アルゴリズムの有用性を検証するために、[13] と正解率と高速性を比較したところ、正解率と高速性共に低下してしましたが、その原因としては、視線モデルの性能の低さが考えられる。

Proposal of Score-Following Reflecting Gaze Information on Cost of DP matching

TERASAKI SHIORI^{1,a)} TAKEGAWA YOSHINARI^{2,b)} HIRATA KEIJI^{2,c)}

1. はじめに

本研究の目的は、従来の打鍵情報による楽譜追跡結果に視線情報を反映させることで、同一の音形の判別や演奏が再開される箇所（演奏再開箇所）の予測を可能とした楽譜追跡システムの構築である。

楽譜追跡は自動伴奏や自動譜めくりの構成技術として多く研究されてきた [1], [2], [6], [8], [9], [11], [15]。従来の楽譜追跡の多くは楽譜追跡情報をもつ演奏の音響信号や MIDI などの打鍵情報をもとに演奏箇所を推定している。そのため、同一の音形の判別や、演奏が一時中断された後に、直前の演奏箇所と全く異なる箇所から演奏を再開され

た場合には楽譜追跡が難しい。このように打鍵情報のみからでは演奏箇所や演奏再開箇所の推定が難しい状況においては楽譜追跡に遅延や誤りが発生する。そこで、我々は打鍵情報とは独立に演奏箇所情報をもつ視線 [12] に着目し、視線情報を打鍵情報を組み合わせた楽譜追跡にシステムを開発した [13]。しかし、サッケードや視線のブレがノイズとなり、演奏者が見ている譜面の範囲をシステムが誤認識し、楽譜追跡精度が下がってしまうという問題があった。また、打鍵情報に依存して演奏箇所の候補を決定するため、打鍵情報から演奏箇所の予測が困難である場合は、演奏箇所の候補に正しい演奏箇所が含まれず、演奏箇所推定ができないといった問題があった。そこで、視線のノイズを除去するために、隠れマルコフモデル（Hidden Markov Model: HMM）によって、視線の動きを予測するモデル（視線モデル）を構築し、視線のノイズ除去を試みる。さらに、視線尤度を算出し、DP マッチングのコストに反映させることで、視線に依存度を高めた楽譜追跡アルゴリズムを実装

¹ 公立はこだて未来大学大学院
Graduate School of Future University Hakodate

² 公立はこだて未来大学
Future University Hakodate

a) g2116028@fun.ac.jp

b) yoshi@fun.ac.jp

c) hirata@fun.ac.jp

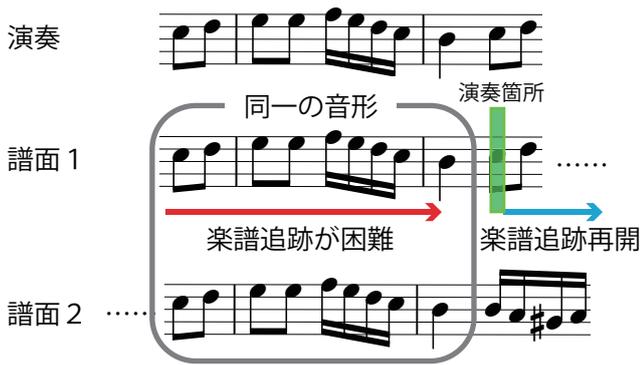


図 1 同一の音形を含む譜面において正しく楽譜追跡が再開されるまでに必要な譜面進行の例

し、打鍵情報から演奏箇所の推定が難しい場合に対応した楽譜追跡を目指す。

2. 関連研究

2.1 自動伴奏システム

Dannenbergs の自動伴奏システムでは演奏と譜面との一致した音の数（ベストマッチ）をコスト関数に設定し、Dynamic Programming（動的計画法、DP）マッチングによって単旋律演奏の楽譜追跡を実現している [2]。ベストマッチを求めていくことで 1 音程度の誤打鍵や弾き飛ばしに対応した楽譜追跡ができる。

一般的に楽譜追跡は HMM によって確率的に演奏箇所の推移を推定することで実現されることが多い [1], [8], [9], [10]。HMM を用いた自動伴奏システムに中村らの Eurydice [8] がある。Eurydice は演奏生成を HMM によってモデル化し、Viterbi アルゴリズムを改良して最尤推定の計算コストを抑えて演奏箇所を探索することで、大幅な跳躍や弾き直しにも対応した高速な楽譜追跡を実現している。また、中村らは同一の音形を含む楽曲において楽譜追跡の誤り率を評価している。演奏の跳躍を含む場合の誤り率について、同一の音形を含まない楽曲と比較すると、同一の音形を含む楽曲に対する誤り率が 3% 増加している。このように、従来の自動伴奏システムは打鍵情報から演奏箇所を探索するため、譜面内に複数同一の音形が出現している場合、譜面に違いが現れるまで、どのフレーズが演奏されたかを即座に特定することが難しい（図 1）。同一の音形を含む楽曲は数多く存在し、練習環境においては譜面の最初から順に演奏されるとは限らないため、この問題を解決することが望ましい。

2.2 演奏と独立に演奏箇所の情報をもつ視線

安価な視線計測装置の出現により、手軽に視線情報が得られるようになり、視線の研究が盛んに行われている。また、演奏中の視線を計測する研究も存在する [4], [5], [7]。Sloboda は視線と演奏箇所の差異（EHS: Eye-Hand Span）

を測定する実験を行い、演奏者が演奏箇所よりも先を視認していることを検証している [12]。このことから、視線の位置座標から逆算すれば、演奏箇所が推定できると考えられる。また、Sloboda は EHS が長い演奏者ほど誤打鍵が少ないことから、熟達度が高いことも述べている。つまり熟達度によって先読みしている音符の数が異なる。また、川崎らは演奏者の楽譜の認識の違いがあるかについて調査するため、演奏中の眼球運動を分析した [3]。この研究では読譜中の譜面上での演奏行為と注視点の位置、停留時間の関連性に注目している。その結果、初級者が初見で譜面を見ながら演奏する場合は、譜面上の音符を逐次的に一つずつ注視するが、練習を重ね、楽譜の構造を理解すると、初級者は音符をまとまりとして認識できるようになることがわかった。従って、ほとんどの演奏者は演奏中に音符をまとまりごとに認識し、譜面箇所を視認しているといえる。

3. 視線を活用した楽譜追跡システム

3.1 システム構成

提案システムの構成を図 2 に示す。本システムは MIDI データ（打鍵位置や打鍵強度）と視線情報（位置座標）を入力とし、楽譜追跡を行う。入力として視線情報が得られている場合は、視線モデルを利用し、視線の生データから演奏者が見ている範囲を推定することで、視線のノイズを除去する（3.2 節）。また、視線モデルの出力確率から算出される視線尤度を反映させた DP マッチングを行い（3.3 節）、演奏箇所を推定する。ただし、視線情報、打鍵情報のどちらか、または両方が検出されない場合は上記の楽譜追跡が難しいため、それぞれ取得している情報から演奏箇所を推定する（3.4 節）以上の過程で推定された演奏箇所を譜面に提示する。

視線追跡装置として、EyeTribe [14] を用いる。EyeTribe は小型かつ据え置き型の視線計測装置であるため、演奏者の演奏を妨げない。また、EyeTribe は 30Hz で視線情報のサンプリングを行う。

3.2 視線モデルの構築

人間は視線を高速に移動させることで視対象を認識している（サッケード）。そのため、サッケードが発生しているときや視線が一瞬視対象から外れたときは、ノイズのためにシステムが視対象を誤認識し、楽譜追跡の性能が低下してしまう。そこで、視線のノイズを含む視線の生データから、HMM によってノイズを除去した視線を予測するモデル（視線モデル）を構築する。演奏者が視対象を視認しているとき、視線は視対象付近に多く出現する。ここで、視線が出現する位置座標の確率分布は明確になっていないため、正規分布に従うとする。従って、正規分布に従う確率密度関数を出力確率とする連続型 HMM によって、ノイズを除去した視線を推定する（図 3）。また、演奏者は

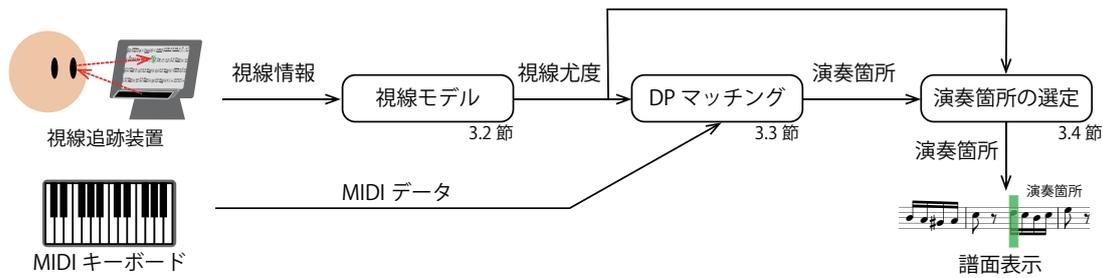


図 2 システム構成

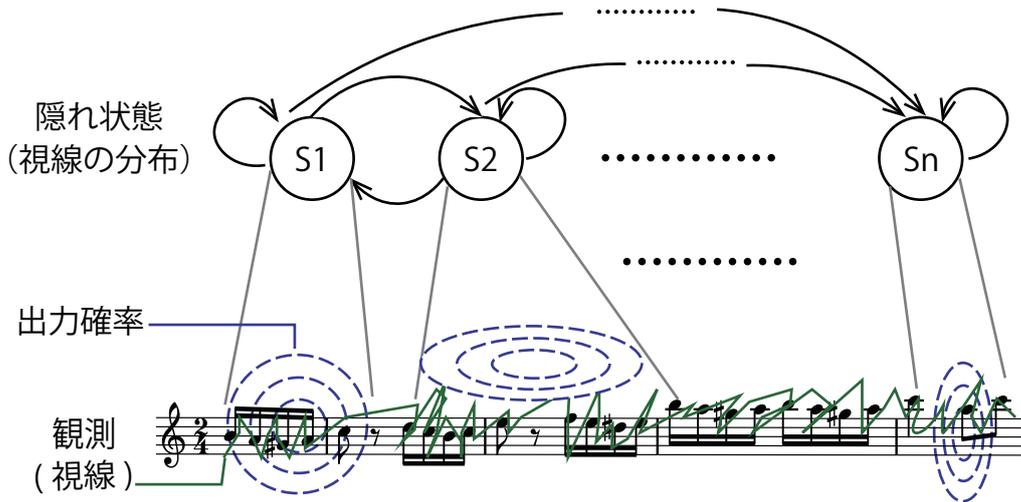


図 3 視線モデル

譜面を読むとき、音符をまとまりごとに認識している [3]. よって、譜面を最小の音符のまとまりごとに視線が分布すると考え、状態は最小の音符のまとまりの数だけ存在するとした。

各状態 s_m から視線 \vec{o}_t が観測される出力確率 $b_m(\vec{o}_t)$ は以下の式で計算される。

$$b_m(\vec{o}_t) = \frac{1}{(\sqrt{2\pi})^D \sqrt{|\Sigma_m|}} e^{-\frac{1}{2}(\vec{o}_t - \vec{\mu}_m)^T \Sigma_m^{-1} (\vec{o}_t - \vec{\mu}_m)} \quad (1)$$

ただし、 D は正規分布の次元数であり、視線モデルは視線の位置座標 (x, y) の二変数を出力変数とするため、 $D = 2$ とする。また、 $\vec{\mu}_m, \Sigma_m$ はそれぞれ状態 s_m から出力される正規分布の平均ベクトルと共分散行列を表す。HMM の状態数を N としたとき、初期状態確率 π , 状態遷移確率 $A = \{a_{ij}\}_{i,j=1}^N$, 出力確率 $B = \{b_i(\cdot)\}_{i=1}^N$ により、HMM のパラメータは $\lambda = (\pi, A, B)$ と表せる。視線の位置座標 $O = \{\vec{o}_1, \vec{o}_2, \dots, \vec{o}_T\}$ が観測されたときの最尤状態 \hat{s}_t は以下の式で計算される。

$$\hat{s}_t = \arg \max_{s_t} P(O, S|\lambda)$$

$$P(O, S|\lambda) = \prod_{t=1}^T a_{s_{t-1}s_t} b_{s_t}(o_t) \quad (2)$$

ここで s_t は状態確率変数である。

3.3 視線情報を反映させた DP マッチングによる楽譜追跡

視線情報を反映させた DP マッチングによって楽譜追跡を行う方法について述べる。DP マッチングによる楽譜追跡は譜面音列と演奏音列の類似度を求め、最も類似度の高い譜面箇所を演奏箇所とする楽譜追跡手法である。従って、1 音程度の誤打鍵や弾き飛ばしであれば、問題なく楽譜追跡できる。通常、DP マッチングでは一箇所楽譜追跡を行うが、大幅な跳躍や弾き戻りや繰り返しには対応できない。そこで、譜面音列から打鍵された音高を全て探索し、全ての譜面箇所候補を作成した。複数の候補でマッチングを行うことで、演奏箇所の候補の信頼度が低い場合に最も信頼度の高い候補を演奏箇所として選択できれば、即座に楽譜追跡の修正が可能である。

例えば、各候補のマッチング結果が大幅に跳躍したとき、演奏者が意図する打鍵位置と近い鍵盤を誤打鍵した可能性があり、マッチング結果の信頼度は低い。そこで、演奏者が演奏時に譜面を先読みするという習性 [12] に基づき、視線情報を DP マッチングに反映すると、マッチングの際に

演奏		視線尤度			演奏		視線尤度		
	H	A	D		H	A	D		
譜面	H	1	1	1	× 0.05	H	1	1	0.05
	A	1	2	2	× 0.05	A	1	2	0.1
	C	1	2	2	× 0.4	C	1	2	0.8
	A	1	2	2	× 0.2	A	1	2	0.4
	H	1	2	2	× 0.1	H	1	2	0.2
	A	1	2	2	× 0.1	A	1	2	0.2
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	D	1	2	3	× 0.05	D	1	2	0.15

大幅な跳躍 → 信頼度が低い

図 4 視線情報を反映させた DP マッチングの例

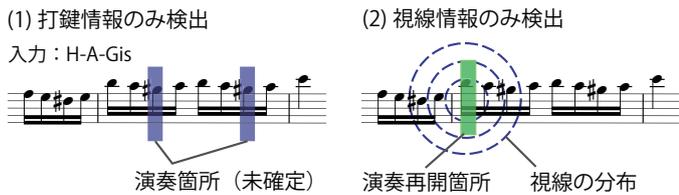


図 5 視線情報、打鍵情報のどちらかが検出されないときのシステムの動き。

スキップした音符中に演奏箇所が存在した場合でも楽譜追跡が可能になる。

DP マッチングのコストに視線情報を反映させるために、視線モデルの出力確率 ((1) 式) をある座標を注視している尤もらしさ (視線尤度) とする。ここで、視線尤度の変化量が大きいと、DP マッチングのベストマッチを求める行列に掛け合わせると、視線による影響が大きくなりすぎてしまい、打鍵情報があまり反映されなくなる恐れがある。そこで、定数 c を (1) 式の exp の指数部に掛け合わせ、視線尤度の変化量を小さくする。

図 4 に視線尤度を反映させた DP マッチングの例を示す。譜面 H - A - C - A - H - A - ... - D に対して、演奏 H - A - D がされたとき、演奏音から DP マッチングを行うと、3 音目の D が演奏されたとき、大幅な跳躍が発生する。このとき、演奏が進行しなければ、マッチングの確かさを判定することは難しい。そこで、ベストマッチを求める行列に視線尤度を掛け合わせる。すると、視線尤度の最も高い (演奏者が演奏している可能性が高い) 3 音目の C の音を演奏箇所として選ぶことができる。このように候補のマッチング結果の信頼度が低い場合には視線尤度を掛け合わせることで打鍵情報のみからは楽譜追跡が難しい場合でも、頑健に楽譜追跡ができる。ただし、候補のマッチング結果が期待通りの場合には、視線尤度を反映させず、マッチング結果を信頼する。

3.4 演奏箇所の選定

視線情報と打鍵情報が共に検出されている場合には 3.3 節で述べたアルゴリズムによって楽譜追跡できる。しか

し、視線情報、打鍵情報のどちらか、または両方が検出されなかった場合には上記アルゴリズムの適用は難しい。そこで、上記アルゴリズムが適用できない場合の処理について考える。それぞれの処理について以下に示す。

- (1) 打鍵はされているが、視線が取得できない場合、手元を見ながら譜面を見ずに演奏している可能性が高い。この場合、視線尤度を無効化すれば、3.3 節のアルゴリズムを適用できる。しかし、視線情報が反映できないため、演奏箇所が 2 箇所以上推定される可能性がある。このとき、演奏箇所を一意に確定するに充たる根拠がないとして、演奏箇所を一つに選定することはしない (図 5 中, (1))。演奏者が再び譜面を視認すれば、即座に演奏箇所を決定できる。
- (2) 打鍵がなく、視線のみを取得している場合、打鍵と打鍵の合間であり、直後に打鍵される可能性が高いため、打鍵情報が入力されるまで待機する。ただし、1 秒以上打鍵されなかった場合は、演奏者が演奏を一時中断している可能性が高い。そこで、演奏再開に備えて、視線によって演奏再開箇所を予測する。視線尤度の最も高い位置にある音符を見ているとして、その音符を演奏再開箇所と予測する (図 5 中, (2))。
- (3) 視線と打鍵が共にない場合、演奏者が休符の間に一瞬手元を確認していたり、演奏以外のことに気を取られている可能性が高い。また、演奏箇所を補正する根拠が得られないため、視線情報が打鍵情報が入力されるまで待機する。

4. 提案アルゴリズムの妥当性の検証

本節では提案アルゴリズムの有用性を段階的に評価する。

4.1 結果

提案した楽譜追跡アルゴリズムの有用性を検証するため、上記アルゴリズムを実装し、正解率と平均推定時間を [13] の楽譜追跡システムと比較する。ここで正解率は全ての打鍵の入力に対し、正しく演奏箇所を推定できた割合とする。また、平均推定時間は誤打鍵が発生してから、正しく演奏箇所を推定するまでの時間とする。平均推定時間が短いほど、高速に演奏箇所の推定ができたといえる。

入力データとして [13] で作成したデータセットを用いる。このデータセットはピアノ初級者以上 7 名にピアノソナタ第 11 番第 3 楽章「トルコ行進曲」の冒頭 16 小節の主旋律を単音で演奏させたときの打鍵情報 (打鍵番号, 打鍵強度) と視線情報 (位置座標) のデータセットである。演奏には跳躍, 繰り返し, 誤打鍵が含まれる (表 1)。トルコ行進曲は冒頭 16 小節中に同一の音形が 3 組存在し、演奏箇所の同定が難しい譜面である。

表 2 に提案アルゴリズムと [13] の楽譜追跡システムの正解率と平均推定時間の比較を示す。このとき、視線モデル

表 1 各演奏に含まれた不確定要素数 (個)

演奏者	跳躍	繰り返し	誤打鍵
A	4	15	13
B	4	15	26
C	4	15	9
D	4	15	21
E	4	15	25
F	4	15	9
G	4	15	11

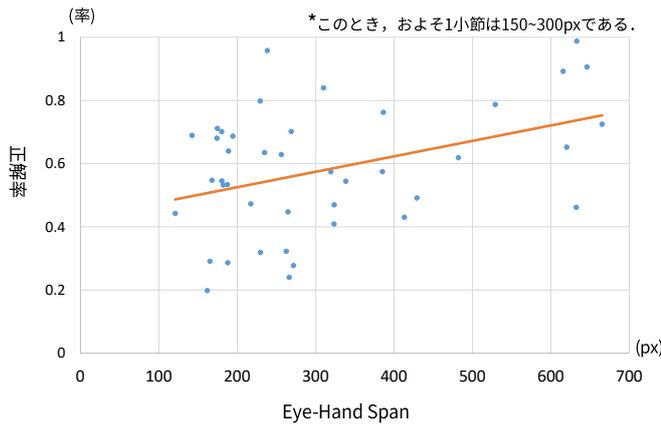


図 6 教師なし学習させた視線モデルを適用した提案アルゴリズムの正解率と EHS の相関



図 7 作為的な視線モデル

のパラメータを決定するために、演奏者が実際に譜面を見ながら演奏している視線データ 40 分を教師なし学習させた。表 2 を見ると、視線認識率の平均は 66.1 % であり、演奏者が演奏中に譜面を認識していることがわかる。[13] と提案アルゴリズムを比較すると、正解率は 62.2 % 低下し (96.8 % から 24.6 %)、平均推定時間は 2.02 秒長くなった (0.48 秒から 2.5 秒)。正解率と高速性ともに著しく低下してしまう結果になった。

ここで、図 6 に正解率と EHS の関係を示す。正解率と演奏箇所と視線の距離の相関を算出すると、これらには正の相関が見られた ($r=0.40$, $p < 0.1$)。これは、熟達した演奏者であるほど EHS が大きいことから [12]、熟達した演奏者の演奏に対しては楽譜追跡の精度がよいことを示している。熟達した演奏者であるほど誤打鍵が少ないため、楽譜追跡の結果として尤もらしい結果であり、提案アルゴリズムの妥当性を示す一つの指標となり得る。

正解率と平均推定時間が低下してしまった原因として、視線モデルの性能が考えられる。提案アルゴリズムは、視

線尤度に楽譜追跡結果が大きく影響されるため、視線モデルの性能が低い場合は楽譜追跡が上手くいかないことが考えられる。そこで、作為的に視線モデルを構築し、視線モデルの性能を考慮しない場合について考える。作為的な視線モデルは演奏者が演奏中に、音符をまとめて見ているという前提に基づき [3]、譜面を最小の音符のまとまりに分割し、それぞれの中心座標を平均値とした正規分布に従う確率密度関数を出力確率として与えた (図 7)。状態遷移確率は人手で決定することが難しかったため、同様に演奏者が実際に譜面を見ながら演奏している視線データ 40 分を教師なし学習させることで決定した。

教師なし学習させたときの視線モデルを適用した提案アルゴリズムと作為的な視線モデルを適用した提案アルゴリズムの楽譜追跡結果を比較すると、全体の正解率が 16.6 % 上昇している (50.2 % から 66.6 %) ことがわかる。また、平均推定時間を見ると、0.41 秒短く (1.35 秒から 0.94 秒) なったことがわかる。つまり、視線モデルが正しく構築されれば、十分に楽譜追跡の向上が見込める。

また、[13] の楽譜追跡アルゴリズムと作為的な視線モデルを適用した提案アルゴリズムの楽譜追跡結果を比較すると、提案アルゴリズムの正解率は 19.1 % 低い (85.7 % から 66.6 %)。また、平均推定時間を見ると、0.6 秒長くなっている。

4.2 考察

[13] と比較すると、正解率と高速性が共に著しく低下してしまうという結果になった。提案アルゴリズムは常に DP マッチングのコストに視線尤度を掛け合わせているため、視線尤度が適切に算出されないとき全体の性能が下がってしまうことが考えられる。楽譜追跡結果は視線尤度に大きく影響される。つまり、視線モデルの性能が低ければ、正しく楽譜追跡ができない。視線モデルの性能が低い要因と解決方法について以下の 2 つが考えられる。

(1) 学習データの少なさ

教師なし学習で汎用的な視線モデルを構築するためには膨大な学習データを要する。しかし、現在は全部で 40 分程度の演奏中の視線データしか持ち合わせておらず、汎用的な視線モデルを構築するためには学習データが不足しているため、視線モデルの性能が低い可能性が高い。そこで、学習のための演奏中の視線データを収集し、視線モデルに学習させることで、安定した停留点が得られる視線モデルの構築を目指す。

(2) 演奏者の熟達度の未考慮

演奏者は熟達度によって EHS が異なる。しかし、現状では各演奏者に合わせた EHS を考慮していないため、EHS が大きい演奏者が誤打鍵を頻発した場合、演奏箇所の視線尤度が小さくなり、楽譜追跡の性能が低下することが考えられる。そこで、視線モデルの学習

表 2 教師なし学習させた視線モデルを適用した提案アルゴリズムと作為的な視線モデルを適用した場合の提案アルゴリズムと [13] 楽譜追跡システムの楽譜追跡の正解率と平均推定時間

演奏者	視線認識率 (%)	正解率 (%)			平均推定時間 (sec)		
		提案 (教師なし学習)	提案 (作為)	[13]	提案 (教師なし学習)	提案 (作為)	[13]
A	73.0	52.9	75.3	90.4	1.57	1.05	0.25
B	72.8	43.9	60.6	83.5	1.13	0.89	0.24
C	38.1	35.6	57.4	74.8	1.59	1.07	0.26
D	70.5	64.4	69.7	87.6	0.78	0.68	0.26
E	83.4	56.8	62.2	85.2	1.38	1.01	0.28
F	68.6	65.6	76.9	87.9	1.02	0.80	0.19
G	86.3	47.0	68.8	92.5	2.01	1.06	0.22
合計	66.1	50.2	66.6	85.7	1.35	0.94	0.24

パラメータとして、熟達度を追加すると、熟達レベルごとに視線モデルが構築され、性能の向上が期待される。

また、作為的な視線モデルを適用したときに、[13] の楽譜追跡アルゴリズムより性能が低い原因として、視線情報と打鍵情報の優先順位の問題が考えられる。現在、打鍵によるマッチング結果の信頼度を候補ごとに決定している。そのため、演奏箇所として選ばれている候補において打鍵の信頼度が高いときでも、他の候補において視線尤度が反映されてしまい、優先順位が逆転してしまう可能性があるため、システムの原因究明を早急に行う必要がある。

5. おわりに

本稿では DP マッチングのコストに視線情報を反映させた楽譜追跡システムを提案した。技術的な課題として、視線のノイズの低減や視線情報を反映させる手法の検討があり、HMM によって視線の動きを予測する視線モデルを構築し、出力確率を視線尤度として DP マッチングのコストに掛け合わせることで解決を試みた。提案アルゴリズムの妥当性を示すために、[13] の楽譜追跡システムと正解率と高速性を比較したところ、正解率と高速性が共に低下してしまった。その原因として、視線モデルの性能が考えられるため、作為的に視線モデルを構築したところ、正解率と高速性の上昇が見られた。そこで、今後は視線モデルの性能を高めるために (1) 膨大な学習データの追加と (2) 熟達度合いを考慮した視線モデルの構築を試みる。

謝辞 本研究は JSPS 科研費 16H01744, 26280089 の助成を受けたものです。

参考文献

[1] Cont, A.: ANTESCOFO: Anticipatory Synchronization and Control of Interactive Parameters in Computer Music, *In Proc. International Computer Music Association (ICMC)*, pp. 33-40 (2008).
 [2] Dannenberg, R.: An on-line algorithm for real-time accompaniment, *In Proc. International Computer Music*

Association (ICMC), pp. 193-198 (1984).
 [3] 川崎智子: ピアノ演奏時における読譜の為の眼球運動, *三重大学教育学部研究紀要教育科学*, Vol. 33, pp. 49-66 (1982).
 [4] Kawase, S.: An Exploratory Study of Gazing Behavior During Live Performance, *In Proc. Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM)*, pp. 227-232 (2009).
 [5] Kobori, S. and Takahashi, K.: Cognitive Processes During Piano and Guitar Performance: An Eye Movement Study, *In Proc. International Conference on Music Perception and Cognition (ICMPC)*, pp. 748-751 (2008).
 [6] 前澤陽, 後藤真孝, 尾形哲也, 奥乃博: 拍長の連続性を考慮した潜在的調波配分法に基づくスコアアライメント手法, *日本音響学会講演論文集*, pp. 1071-1074 (2011).
 [7] 中平勝子: 意識—無意識行動変容プロセスモデルと学習者分類—視行動に基づく技能レベル判定—, *日本人間工学会論文誌*, Vol. 50, pp. 90-91 (2014).
 [8] Nakamura, E., Cuvillier, P., Cont, A., Ono, N. and Sagayama, S.: Autoregressive Hidden Semi-Markov Model of Symbolic Music Performance for Score Following, *In Proc. International Society for Music Information Retrieval Conference (ISMIR)*, pp. 392-398 (2015).
 [9] 中村友彦, 中村栄太, 嵯峨山茂樹: 誤り・任意の弾き直し・弾き飛ばしを含む演奏音響信号への高速な楽譜追跡, *情報処理学会研究報告音楽情報科学*, vol. 2013-MUS-99, No. 40, pp. 1-5 (2013).
 [10] Orio, N., and Dechelle, F.: Score Following Using Spectral Analysis and Hidden Markov Models, *Proc. International Computer Music Association (ICMC)*, pp. 1708-1710 (2001).
 [11] Raphael, C., Music Plus One: A System for Expressive and Flexible Musical Accompaniment, *In Proc. International Computer Music Association (ICMC)*, pp. 159-162 (2001).
 [12] Sloboda, J.: The Eye - Hand Span - An Approach to the Study of Sight Reading, *Psychology of Music*, 2(2), . 4-10 (1974).
 [13] 寺崎葉里, 竹川佳成, 平田圭二: 視線情報を活用した楽譜追跡システムの構築, *研究報告音楽情報科学*, Vol. 2016-MUS-112, No. 11, pp. 1-8, (2016).
 [14] The Eye Tribe, <https://theyetribe.com/> (参照 2016-02-06).
 [15] Vecoe, B.: The Synthetic Performer in the Context of Live Performance, *In Proc. International Computer Music Association (ICMC)*, pp. 199-200 (1984).