

研究速報

5 指の指先の同時追跡に基づくピアノ運指認識手法

有賀 治樹[†]西山 乗[†]橋本 学^{†a)} (正員)長田 典子^{††} (正員)

Piano Fingering Recognition by Simultaneous Tracking of Five Fingertips

Haruki ARUGA[†], Jo NISHIYAMA[†], *Nonmembers*,
Manabu HASHIMOTO^{†a)}, and Noriko NAGATA^{††}, *Members*[†] 中京大学大学院情報科学研究科, 名古屋市Graduate School of Information Science and Technology,
Chukyo University, Nagoya-shi, 466-8666 Japan^{††} 関西学院大学大学院理工学研究科, 三田市Graduate School of Science and Engineering, Kwansai
Gakuin University, Sanda-shi, 669-1337 Japan

a) E-mail: mana@isl.sist.chukyo-u.ac.jp

DOI:10.14923/transinfj.2014JDL8020

あらまし 複数の指先の追跡を最適経路決定問題として扱い、解析的 DPT により得られた追跡経路から運指を認識する手法を提案する。複数物体の相対的な位置関係の連続性を評価することにより、特徴が似た物体を同時に追跡できる。実験により、運指認識成功率 98% を確認した。

キーワード DP トラッキング, 複数物体追跡, ワイルドカード, 運指認識

1. まえがき

初心者のピアノ演奏の学習をサポートするために運指の自動認識が望まれているが、従来の運指認識手法 [1], [2] では、指先の密着や、打鍵指のオクルージョンによる誤認識が問題であった。そこで、指先の追跡を動画全フレームにおける指先位置の最適経路決定問題と捉え、大局的な最適解を求めることによって、指先の密着やオクルージョンに頑健な指先追跡を行い、追跡された全ての指先の経路から運指を認識する手法を提案する。大局的な最適化に基づく追跡手法として、解析的解法によって高精度かつ高速に追跡する手法 (解析的 DP トラッキング; DPT) [3] がある。この手法は、テンプレートとの類似度とフレーム間の移動量の連続性を評価する目的関数を用いて、画像上の単一物体を追跡する手法である。しかしながら、それぞれの指先を単独で追跡する場合には、指の種類の違いによる画像特徴の差が小さいことに起因して、他の指先との誤対応が起こり得る。

そこで本研究では、指先同士の相対的な位置関係の連続性を評価する項を解析的 DPT の目的関数に加えることにより、解析的 DPT をベースとして五つの指先を同時に追跡し、追跡経路から運指を認識する。こ

のとき、事前に指先検出位置を探索空間とした DP トラッキング (離散的 DPT) を行い、解析的 DPT のための粗い追跡経路を算出する。また、指先位置を固定しないワイルドカードを探索空間に設定することによって指先の未検出を考慮し、オクルージョンに頑健な安定した追跡を可能とする。

2. 提案アルゴリズム

提案手法のアルゴリズムを図 1 に示す。まず、フレームごとに指先を検出し、候補指先位置に設定する。次に、この候補指先位置をもとに離散的 DPT と解析的 DPT を用いて五つの指先を同時に追跡し、指先追跡経路を算出する。最後に、この追跡経路と打鍵された鍵盤位置を比較することによって運指を判定する。

追跡の 1 段階目の処理である離散的 DPT では、想定される候補指先位置と全指先の対応付けの中から最適な対応付けを一括処理により算出する。その後、2 段階目の処理である解析的 DPT では、画像全体を探索空間として最適な追跡経路を解析的に算出する。

2.1 離散的 DPT を用いた指先対応付け

解析的 DPT の前処理として、候補指先位置を探索空間とした全指先との対応付けを離散的 DPT により行う。第 t フレームの追跡パラメータを $\mathbf{w}_t = (\mathbf{w}_{t,1}, \dots, \mathbf{w}_{t,M})^T$, $\mathbf{w}_{t,m} = (x_{t,m}, y_{t,m})^T$ と定義して、目的関数 (1) を最小化する最適経路 ($\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_T$) を求める。ただし、 M は指先の数、 m は指の種類を示しており、 T は動画像のフレーム数である。また、 λ_1 と λ_2 は各項の重みである。

$$F(\mathbf{w}_1, \dots, \mathbf{w}_T) = \sum_{t=1}^T \sum_{m=1}^M d_{t,m}(\mathbf{w}_{t,m}) + \lambda_1 \sum_{t=2}^T \sum_{m=1}^M (\mathbf{w}_{t,m} - \mathbf{w}_{t-1,m})^2 + \lambda_2 \sum_{t=2}^T \sum_{m=2}^M \{(\mathbf{w}_{t,m} - \mathbf{w}_{t,m-1}) - (\mathbf{w}_{t-1,m} - \mathbf{w}_{t-1,m-1})\}^2 \quad (1)$$

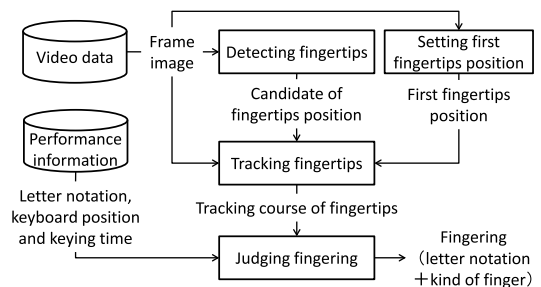


図 1 提案アルゴリズムのブロック図

Fig. 1 The block diagram of proposed algorithm.

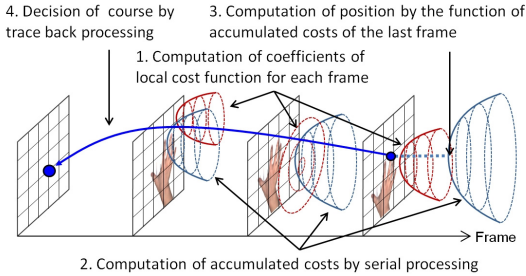


図2 最適経路の決定手順

Fig. 2 The procedure for determining the optimal path.

第1項は、第 t フレームにおいて m の指先が $\mathbf{w}_{t,m}$ に存在することを仮定した場合の局所コスト $d_{t,m}(\mathbf{w}_{t,m})$ を示し、本研究ではChamfer Matchingを用いる。第2項、第3項はそれぞれ、位置の連続性コストと相対位置の連続性コストを意味する。このとき、候補指先位置の中に指先位置を同定しないワイルドカードを設定することによって、セルフオクルージョンによる指先の未検出に対応する。目的関数の計算時には他の指先位置から推定した指先位置を用いて評価する。目的関数を用いて最終フレームまで累積コストを計算し、累積コストが最小の位置からバックトレースすることにより、最適経路を決定する。

2.2 解析的DPTを用いた高精度指先追跡

離散的DPTにより得られた追跡経路を中心として局所コストを二次関数で表現することにより、近似的に全探索空間で追跡することができる。図2に解析的DPマッチングの処理手順を示す。各フレームの局所コスト関数と累積コスト関数の係数を算出後、最終フレームの累積コスト関数の最小の位置からバックトレースすることにより、最適経路を決定する。

3. 実験結果と考察

実験には初心者向けの演奏楽曲としてMenuet (J.S.Bach)の一部を演奏する約32秒間の動画像(30fps, 960枚, 64打鍵)を使用した。実験環境を図3に示す。実験には、電子ピアノとして(a)に示したカシオ計算機株式会社のPX-750BK, RGBカメラとして(b)に示したThe Imaging Source社のDFK22AUC03を用いた。実験に用いた動画像は、鍵盤から高さ0.8m, 鍵盤の奥から手前に向けて約10度の角度で設置したカメラを用いて取得した。また、約3オクターブの鍵盤を計測可能であり、画像サイズは744x480である。(c)に実際に取得した画像の例を示

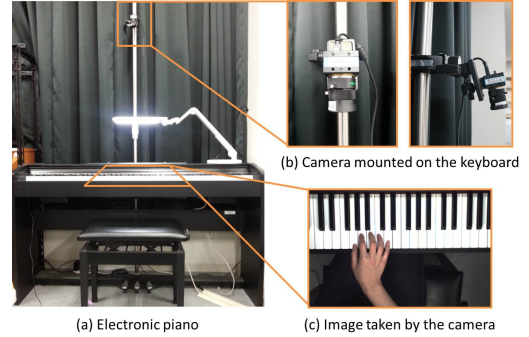


図3 実験環境

Fig. 3 The environment of the experiment.

表1 追跡成功率と処理時間

Table 1 Tracking success rate and processing time.

提案手法	追跡成功率 [%]					処理時間 [sec]
	親指	人差指	中指	薬指	小指	
Mean Shift 法	55.0	29.9	1.7	0.9	6.3	18.8
離散的 DPT	49.1	97.5	97.1	96.6	97.7	87.6
提案手法	83.1	97.3	96.8	97.2	98.9	94.6

(CPU: Intel®CORE™i7-2.93GHz, メモリ: 8GB)

す。演奏楽曲は、一般的には初心者向けの演奏楽曲とされており、指先のセルフオクルージョンが多数発生しているが、ピアノの練習曲として一般的に使用される楽曲である。また、本実験では目的関数(1)の λ_1 を1, λ_2 を7とした。

まず、指先追跡の比較実験を示す。比較手法は色ヒストグラムを用いたMean Shift法[4]と離散的DPTのみを用いた手法、提案手法とする。結果を表1に示す。追跡成功率は追跡結果と正解位置との距離がしきい値(15画素)以下であった割合とした。

Mean Shift法では、指同士の密着により他の指への誤対応が発生したため、認識率が低下していた。特に指同士の密着が頻繁に起きていた中指、薬指、小指の認識率が低下していた。これに対して、提案手法は誤対応することなく、安定した追跡が可能であった。指ぐりによって、親指が遮蔽されながら大きく移動しているため、離散的DPTと提案手法においては、親指が他の指に比べて追跡成功率が低くなったと考えられる。しかし、提案手法は離散的DPTに比べて、親指に対して追跡成功率の向上が確認できた。

次に、運指認識性能を表2に示す。運指認識成功率は、認識結果の運指と実際の演奏で用いた運指が一致した割合とした。鍵盤の押下の有無は、電子ピアノから得られるMIDI信号の音名、打鍵開始時刻、打鍵終了時刻の情報をもとに判断する。演奏曲に対する提案

表 2 運指認識成功率と処理時間

Table 2 Fingering recognition rate and processing time.

追跡手法	認識成功率 [%]	処理時間 [sec]
Mean Shift 法	40	14.7
離散的 DPT	95	49.7
提案手法	98	52.4

手法の認識成功率は 98%，処理時間 52.4 秒であった。運指は、実際に打鍵した指先の追跡結果が打鍵された鍵盤の範囲内（横幅：約 30 画素，縦幅：約 220 画素）であれば正しく認識できるため，表 1 の追跡成功率に比べて運指認識成功率が高くなったと考えられる。本手法は，演奏が終了した時点から処理を始め，例えば 30 秒程度の演奏の場合，演奏後 50 秒程度で結果を出力できる。したがって，演奏を記録した動画像に合わせて運指認識結果を演奏者に提示することによって，演奏の支援が十分に可能であると考えられる。

4. む す び

解析的 DPT をベースとして，新たに複数物体の相対的な位置関係を考慮する項を目的関数に加えることにより，複数の指先を同時に追跡し，追跡経路から運

指を認識する手法を提案した。指先のセルフオクルージョンを含む，実演奏データに対する実験により本手法の有用性を確認した。

謝辞 本研究の一部は科研費 B (24300088) の助成を受けて遂行されたものである。

文 献

- [1] 子安大士, 木村慎二, 前川 仁, “ピアノ演奏動作解析のための 3 次元手指追跡,” 情報科学技術フォーラム講演論文集, vol.10, no.3, pp.171–172, Sept. 2011.
- [2] 有賀治樹, 岡 明也, 橋本 学, 長田典子, “連続デプス画像解析に基づく仮説検証型ピアノ運指認識手法,” 情処学音楽情報科学研報, vol.2013–MUS–100, no.11, pp.1–6, Aug. 2013.
- [3] 藤村一行, 内田誠一, “大局的最適化に基づくトラッキング-DP トラッキング,” 信学論 (D), vol.J92-D, no.8, pp.1279–1288, Aug. 2009.
- [4] D. Comaniciu, V. Ramesh, and P. Meer, “Real-time tracking of non-rigid objects using mean shift,” IEEE Computer Vision and Pattern Recognition, vol.2, pp.142–149, Hilton Head Island, United States of America, June 2000.

(平成 26 年 5 月 18 日受付, 9 月 30 日再受付,
10 月 31 日早期公開)