

ピアノ演奏時の手指画像を用いた接鍵検出法

國澤佳代[†] 浜中雅俊[‡] 星野聖[‡]

筑波大学大学院システム情報工学研究科[†] 筑波大学システム情報系[‡]

1. はじめに

本稿では、ピアノ演奏時の手指および鍵盤画像から得られる高次局所自己相関特徴量から鍵に指が接触したタイミング(以下、接鍵タイミングと呼ぶ)の検出を行う手法について述べる。

ピアノは鍵盤を叩くと決まった音高で音を出すという楽器としては単純な構造をもつ。しかし、演奏する際には滑らかかつ正確に鍵盤を弾くことや、左右の手指で別々の動作をすること、さらに抑揚などを表現するといったことが必要となるため、挫折してしまう初学者が多い。これらの理由から、ピアノ演奏初心者のための運指支援システムの研究がなされているが、抑揚表現などに代表されるアーティキュレーションの支援に関する研究は進んでいない[1][2]。

また、ピアノ演奏初心者のアーティキュレーション支援のためには、打鍵時、特に指と鍵が接触した瞬間の手指のフォームが重要である。そこで、手指に焦点を当てたピアノ演奏画像から接鍵タイミングの検出を行う。通常の電子ピアノや MIDI キーボードでは、鍵が押し込まれた時に打鍵信号を検出するため、接鍵タイミングの検出はできない。

一方、画像からの認識手法として、kinect の利用や、オプティカルフロー処理を用いたトラッキング手法などがあるが、これらの手法ではオクルージョンが発生してしまった場合、手指の認識精度が落ちてしまうため、接鍵タイミングにおける手指の検出は困難である。

画像特徴量である高次局所自己相関特徴量から手指認識を行うといった研究もなされているが、ピアノ演奏時の手指は形状が似通っており、通常手指認識で用いられるような、動作が大きく変化するものではない[3]。

しかし、オクルージョンによって指先の検出が困難な場合であっても、手指の微細な動作か

ら高次局所自己相関特徴量を抽出することで、接鍵タイミングが検出できる可能性がある。そこで我々は、取得動画のフレーム間の差分画像に着目し、高次局所自己相関特徴量を抽出、その成分を解析した後、SVM を用いた分類器による接鍵タイミングの推定を行った。

2. 接鍵検出システム設計

ピアノ演奏時の手指および鍵盤画像から接鍵タイミングを得るシステムの検討を行う。

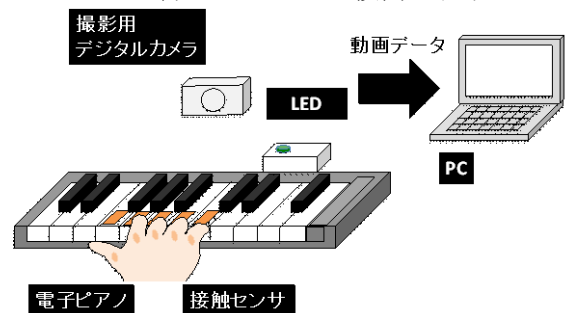


図 1: システム概要

2.1 システム構成

提案する手指画像からの接鍵検出システムの概要図を図 1 に示す。鍵盤には指の接触を感知する接触センサを取り付けてあり、接触の際に LED が点灯する。カメラは LED、ピアノ演奏時の手指、および周辺環境を撮影し、演奏者はビデオ録画が始まると演奏を開始する。この演奏動画は PC で処理する。ここで LED の点滅は評価における正解データの正誤判定に用いるため、特徴量を抽出する際の動画は演奏時に打鍵される鍵盤、奏者の手指全体が映っている箇所をトリミングして使用する。

2.2 画像処理法とシステム

ピアノ演奏時の画像情報から接鍵判定を行うためのシステムの流れを検討する。

画像処理の流れを図 2 に示す。動画に対してフレーム間差分処理を行い、二値化を行う。フレーム間差分処理を行うことによってフレーム間で動きのある部分のみが検出されるため、指の

“Method of keystroke detection using hand image in piano performance”,[†]Kunisawa Kayo, University of Tsukuba, Graduate School of Systems and Information Engineering, [‡]Hamanaka Masatoshi, Hoshino Kiyoshi, University of Tsukuba, Faculty of Engineering, Information

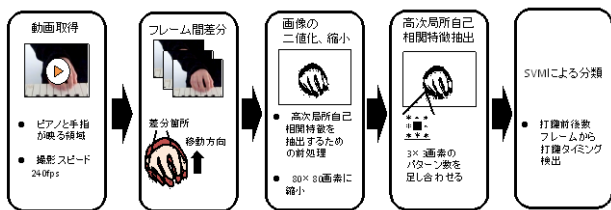


図 2：画像処理の流れ

みの移動などの微細な移動に対しても変化が検出され、特徴量の差異が観測可能となる。その後、二値化処理および画像の縮小処理を行うことにより、高次局所自己相関関数の特徴量を検出する。さらに、接鍵タイミングとその前後 1 フレームにおける高次局所自己相関特徴量を入力として SVM を学習し、接鍵タイミングの検出を行う。

2.3 画像特徴量

本稿では、画像特徴量として、手指の形状推定で高い性能を示している高次局所自己相関特徴[3][4]を採用する。高次局所自己相関特徴は、 N 次自己相関関数に基づき、画像から抽出される特徴において、1)位置に関する不変性、2)画面に関する加法性、3)学習による適応性を満たしている。自己相関関数は平行移動に対して不変であるということから、高次自己相関関数は、変位方向 (a_1, a_2, \dots, a_N) に対して、

$$x^n(a_1, a_2, \dots, a_N) = \int f(r)f(r+a_1)\dots f(r+a_N) dr \quad (1)$$

で定義される。画像の相関特徴に適應する際には、変位方向は参照点 r を局所領域に限定する必要がある。局所領域は、 3×3 画素とし、次元数 n は 2 とする。平行移動による等価な特徴を除くと、二値画像において局所画像は 25 個のパターンとなる。(図 2) これらの特徴から画像中でどの成分が多いかを判定する。



図 3: 高次局所自己相関特徴パターン (No. 6 以降は省略)

2.4 分類器

差分画像から得た高次局所自己相関特徴量を 3 フレーム 1 組とし、サポートベクターマシン(以下、SVM と呼ぶ)によって分類する。

3. 実装

2 章で述べた接鍵検出システムの実装を行った。機材は、Macbook Air(1.8GHz, Apple 製)、電子ピアノ (PSR-E303, YAMAHA 製)、LED 点灯用に Arduino UNO(16MHz)、デジタルカメラ (EX-ZR400, CASIO 製)を使用した。動画は室内で撮影を行い、カメラをキーボードに固定し撮影を

行った。撮影スピードは 240fps、高次局所自己相関特徴量を抽出する画像は 80×80 画素とした。

4. 実験

ピアノ演奏時の手指やその他の部分を撮影した動画から SVM を用いて接鍵検出を行った。評価は、以下の手順に従って行う。撮影画像 7793 枚の教師データを学習させ、テストデータ 1096 枚から正解率を求める。

課題曲：童謡のきらきら星の冒頭 4 小節、バイエルから 38 番、57 番、64 番の冒頭 8 小節の 4 曲のメロディ部分を演奏してもらった。

被験者：各曲それぞれ 3 名ずつ実験を行った。

結果：高次局所自己相関特徴の結果を入力として SVM による接鍵タイミングの識別を行った。その結果、97.91%の正解率となり、本手法の有用性が確認できた。

5. まとめと展望

本稿では、ピアノ演奏時の手指および鍵盤画像から得られる高次局所自己相関特徴量から鍵に指が接触したタイミングの検出を行う手法について述べた。具体的には、取得動画のフレーム間の差分画像に着目し、高次局所自己相関特徴量を抽出、成分を解析後、SVM を用いた分類器による接鍵タイミングの推定を行った。

その結果、97.91%の正解率となり、本手法の有用性が確認できた。今後さらに多くの曲について実験を行い、性能を評価したい。

謝辞

本研究の一部は、総務省戦略的情報通信研究開発推進制度 (SCOPE) の援助により行われた。

参考文献

[1] 竹川佳成, 寺田 努, 塚本昌彦: リズム学習を考慮したピアノ演奏学習支援システムの構築, 情報処理学会インタラクシオン 2012, pp. 73-80
 [2] 岡明也, 有賀治樹, 杉山健太郎, 橋本学, 長田典子: マーカレス運指認識と音列照合によるピアノ演奏スキル評価システムの提案, 情報処理学会研究報告 (音楽情報処理学会 2013-MUS-100), No.10
 [3] 富田元将, 星野聖: 自由な前腕動作を許容する手指形状システム, 電子情報通信学会技術研究報告. EID, 電子ディスプレイ 109(223), pp.9-12, 2009-10-01
 [4] 栗田多喜夫: 柔らかな情報処理のための統計的手法の応用に関する研究, 電子技術総合研究所研究報告, 第 957 号, (1993.11)