

距離画像のガウスフィッティングによる指位置推定に基づく マーカレスピアノ運指認識手法

内田 駿[†] 岡 明也[†] 橋本 学[‡]

中京大学 情報理工学部[†] 中京大学大学院 情報科学研究科[‡]

{uchida, mana}@isl.sist.chukyo-u.ac.jp

1. はじめに

正しい運指で演奏することはピアノ初心者の演奏スキル向上のために重要である。運指とは、打鍵した指の種類と打鍵された音名の組み合わせ情報である。運指を目視で判断することは演奏者にとって困難であることから自動的に認識し、運指の間違いを指摘するシステムが望まれている。

従来のピアノ運指認識の研究例として、手領域を抽出するために背景差分を用いた手法[1]がある。しかし、手指と白鍵のコントラストが十分でないため、安定した手領域の抽出は困難である。また、エッジ画像に対し主曲率を求めることで指先を検出する手法[2]では、画像上で各指が密着する状態において安定的に指先を検出することが困難である。さらに、学習ベースによる手法[3]では、各指が密着している状態の画像を学習すれば認識可能であるが、学習に必要なデータ量が膨大になるという問題がある。

本稿では、1枚の距離画像から各指の位置を推定することで打鍵指を認識する手法を提案する。打鍵指については、距離画像から得られる指領域の濃度プロファイルに複数のガウス分布を当てはめることで、各指の位置を推定する。当てはめには、EMアルゴリズムを用いて、ガウス分布のパラメータを推定して各指の位置推定を行う。なお、打鍵された音名信号は、電子ピアノのMIDI情報として取得する。これらを組み合わせることで運指情報を生成し、出力する。

2. 提案手法

2.1. 概要

まず、距離画像を取得し手領域を抽出する。次に、抽出した手領域から手首位置を中心とした楕円上に指候補領域の検出を行う。そして、指候補領域の濃度プロファイルに混合ガウス分布を当てはめて各指の位置を推定し、打鍵され

た鍵盤の位置に存在する指を打鍵指とする。推定した指位置とキーボードから得られる音名信号(MIDI信号)を用いて運指認識する。図1に提案手法のフローチャートを示す。



図1. 提案手法のフローチャート

2.2. 手領域の抽出

打鍵時の画像を取得するために、一般的なMIDI仕様のキーボードを用いる。MIDI信号を用いることにより打鍵イベント、離鍵イベントの信号が取得可能である。打鍵イベントの信号を取得した際に撮影することで、打鍵時の画像を取得できる。打鍵時に撮影した画像に対し、鍵盤のみの画像を背景画像として背景差分を施し手領域を抽出する。距離情報を用いることにより安定的に手領域を抽出することができる。

2.3. 指候補領域の検出

手首位置を中心とした楕円の長軸を指先まで伸ばすことにより、指候補領域の検出を行う。図2に楕円の設定を示す。短軸aの長さは手領域の横幅をもとに求める。長軸bの長さを求める際に、手領域の面積を楕円の面積、短軸を手領域の横幅として求める。また、画像における手の方向の変化により指候補領域が安定的に検出できない場合があるため、手の方向を求める必要がある。手の方向は、楕円の長軸を、手首位置と予め求めておいた手領域の重心を結ぶ直線に合わせて長軸を回転させることで求める。以上より、画像における手の方向による変化に対応する。

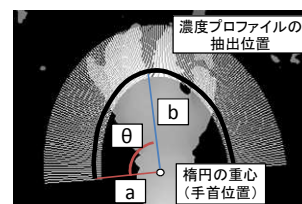


図2. 手領域における楕円の設定

A method for marker-less piano fingering recognition based on estimation using gaussian fitting for depth image

[†]Suguru Uchida, Akiya Oka, School of Information Science and Technology, Chukyo University.

[‡]Manabu Hashimoto, Graduate School of Information Science, Chukyo University.

2.4. 混合ガウス分布による各指の位置推定

検出した指候補領域の濃度プロファイルに対してEMアルゴリズムを用いて式(1)で示す混合ガウス分布を当てはめる. 式(1)のGMMは混合ガウス分布を示し, G はガウス分布を表す. 図2の曲線が濃度プロファイルの抽出位置である. 同図の θ は, 方位角を示す. μ_i は指の中心位置, σ_i は指の太さに相当する. なお, 5本の指が画像上に写ることを前提とし, ガウスの数は5に固定する.

$$GMM(x) = \sum_{i=1}^5 \omega_i G_i(x; \mu_i, \sigma_i^2) \quad (1)$$

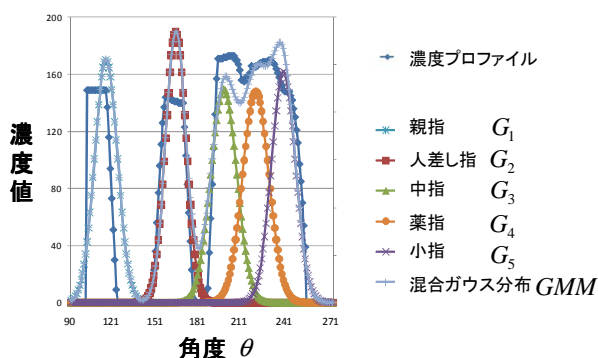


図3. 濃度プロファイルに対するEMアルゴリズムのパラメータ推定結果

図3の $G_1 \sim G_5$ は各々の親指から小指の位置を示しており, 画像上で各指が密着している状態において各指の位置が推定可能であることがわかる. さらに, 楕円の長軸を指先まで伸ばす度に各指の位置を同様に推定する. 楕円の長軸を伸ばしても各指の位置変化が少ないことから, 前回の楕円の濃度プロファイルにおける推定パラメータを初期値とする. 図4に推定された各指の位置を円を用いて示す.

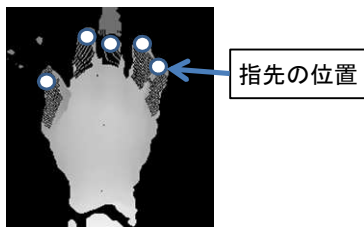


図4. 各指位置推定の結果

3. 提案手法による認識結果と考察

本研究ではレンジセンサとして, Microsoft®のkinectを用いた. kinectは, 鍵盤の直上70cmに設置された.

3.1. 各指が密着している状態の画像での比較実験と実画像における正答率評価実験

画像上で各指が密着している状態において従来法との比較実験結果を表1に示す. 従来法で用いた画像は, レンジセンサとしてTOFカメラを用いて取得された.

表1. 比較実験結果

	画像上での密着状態画像	認識結果
従来法 [3]		
提案手法		

従来法では各指が密着している状態で誤認識したが, 提案手法においては認識することができた. また, 実画像100枚における正答率を表2に示す.

表2. 実画像における正答率

	正答率
各指が密着している状態の画像	44/50(88%)
各指が密着していない状態の画像	45/50(90%)

※打鍵指と音名の両方が正しいときに正答とした.

実験結果より本手法は, 画像上で各指が密着している状態に対しても有効である. 誤認識の主因は, 指候補領域が安定的に検出されず, 手全体の領域まで検出したためである.

3.2. 提案手法におけるデータ量に関する考察

従来法では, 鍵盤全範囲をカバーするために学習画像のメモリ量は約8GB必要であるが, 本手法は不要である.

4. おわりに

混合ガウスモデルを用いて, 実際に撮影した距離画像より指位置の推定を行った.

参考文献

- [1]Dmitry O, et al. "Detection and tracking of pianist hands and fingers". Proc.of CRV, 2006
- [2]子安, 木村, 前川, "ピアノ演奏動作解析のための3次元手指追跡", FIT, 2011
- [3]岡, 橋本, "音名信号を用いた学習データ選択に基づく距離画像ベースのマーカレスピアノ運指認識手法の提案", SIGMUS, 2011