

# 単眼カメラによるハンドジェスチャー認識のための高速手指検出

根本 祐也<sup>†</sup> 中島 克人<sup>†</sup>

東京電機大学大学院未来科学研究科<sup>†</sup>

## 1. はじめに

単眼カメラによるハンドジェスチャー認識は、安価にシステムを構成できるが、手の認識速度と精度に課題がある。特に、実環境では肌色情報などで抽出した手指の候補領域にノイズが含まれるため、ノイズ除去のための前処理に時間をかけるか形状認識の精度を犠牲にするかの選択に迫られる事が多い。

本稿では、複雑背景下での実時間ハンドジェスチャー認識システムを目的としたノイズに強い高速な指先検出手法を提案すると共にその評価結果を報告する。

## 2. 提案手法の概要

本研究ではハンドジェスチャー認識を、(1) 手の位置と腕の方向の検出、(2) 検出位置での手形状識別、(3) 手形状の遷移によるジェスチャー認識、の3段階で行う[1]。これにより、実時間処理のための高速化と認識の精度向上を図る。本稿では、(1)の結果を元にした、(2)を目的とする高速な指先検出手法を提案する

### 2.1 検出位置での手形状識別

指先の検出は手形状を認識するための重要な要素である本研究では、物体検出手法の一つである RRF[2]を参考にした RRD(Radial Reach Detector)を用いて、(1)で検出した手領域の中で尖った形状の領域の検出を行う。

#### 2.1.1 手領域の検出

指先の探索領域は画像全体ではなく、予め探索領域の絞り込みを行い、その領域の周辺でのみ行うのが効率上望ましい。そこで、本手法では、指先検出の前処理として手の候補領域の検出を行う。

手の候補領域の検出には、(1)に関する提案である高速な手検出手法[1]を用いる。

[1]では、ジェスチャーが動きを伴う肌色領域である事に着目した前景抽出手法と、2重矩形を用いた前景画素の比率計測により、手領域を検出する。後者では、内側矩形内の前景画素数が一定以上であることと、その外周、即ち、外側矩形と内側矩形の間の領域の前景画素数が一部にのみ分布していることによって、手の位置と手首の方向を検出する手法である(図1)。内側矩形および外周を分割したそれぞれの小矩形での前景画素数は積分画像により求めるため、非常に高速である。そのため、(2)における指先探索や手形状識別の領域を絞り込む手法として有用である。

なお、[1]による手領域検出では、手形状の違いに基づく手領域の見えの大きさの影響を受けるため、指先の位

置までは正確に求められない。そこで本研究では、検出結果における2重矩形の内側領域の幅が掌の幅に近い事に着目し、内側領域の矩形幅の3倍に当たる領域(図2参照)を指先探索範囲として定めた。

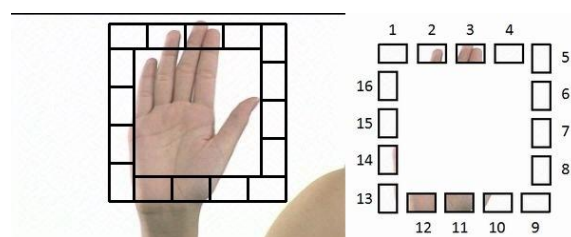


図1. 2重矩形による手領域の検出



図2. 指先探索範囲の設定

#### 2.1.2 指先検出

RRDでは、検出した手領域画素とそれ以外を表す2値画像をベースとする。手領域の画素それぞれを探索基点とし、基点から8方向に長さ $r$ の探索線を伸ばす。各探索線上では、探索線の長さに対する平均の手領域画素数を求め、閾値と比較する。平均画素数がある閾値(例えば80%)以上であればその探索線方向は手領域の内側方向であり、平均画素数が閾値を下回れば、その探索線方向は手領域の外側方向であると推測する。

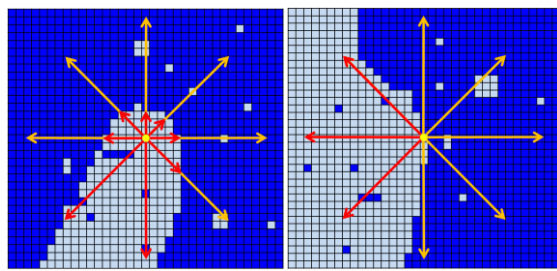
探索線上で手領域画素の連続性を検査するのではなく、平均画素数を検査することにより、ノイズ除去などの前処理を不要にでき、[3]のようなテンプレートマッチングによる指先検出よりも計算量を削減出来るため、高速でロバストな検出が可能となる。

なお、探索基点から見て、閾値以上となる探索線が一方または隣接する二方向であった場合、探索基点は手領域内の尖った形状の領域周辺であると推定し、その探索基点を指の先端の候補画素とする。隣接する二方向を許容するのは指向きの変化に対応する為である(図3参照)。これにより、16方向の指の向きを判定できる事になる。

Fast Hand and Finger Detection for Hand Gesture Recognition using a Monocular Camera

<sup>†</sup>Yuya NEMOTO, <sup>†</sup>Katsuto NAKAJIMA

<sup>†</sup>Graduate School of Science and Technology for Future Life, Tokyo Denki University

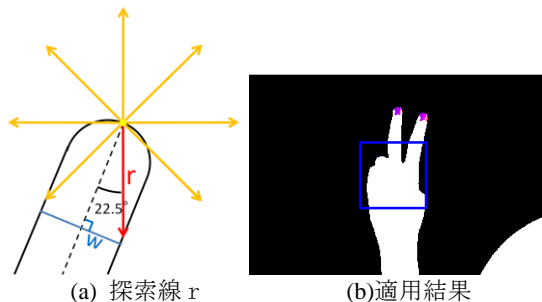


(a)指先候補 (b)非指先候補  
図3. 指先候補画素検出

### 2.1.3 閾値の設定

RRDにおける探索線の長さ  $r$  は指の傾きを考慮して、指の幅  $w$  に対して  $r=(w/2)/\sin 22.5^\circ$  としたい.  $w$  は直接求まらないが、手領域検出時の二重矩形の内側矩形の幅から推定できると考え、内側矩形に対する指の幅の比率の調査を行った. 調査対象には手話の指文字 106 種類の画像[4]に対して手領域検出を実行した結果画像(図2(b)など)を使用した.

調査の結果、二重矩形の内側の矩形幅に対して、指先から指の幅は 12%である事が分かった. そのため、今回は、 $r=(w/2)/\sin 22.5^\circ = ((\text{内側矩形幅} \times 0.12)/2)/0.383 = \text{内側矩形幅} \times 16\%$  とした.



(a) 探索線  $r$  (b)適用結果  
図4. 探索線の設定

## 3. 実験

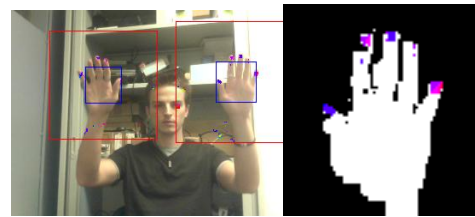
### 3.1 実験とその環境

Spruyt らの実環境手検出のためのビデオデータセット[5]での検出実験と処理速度の計測を行った. 走査矩形のサイズは、入力画像の短辺長から  $20 \times 20$  画素まで 0.9 倍ずつ縮小する. なお、フレーム内で一度検出に成功した場合、3 回縮小した後に探索を終了する. 今回の実験に用いる検出器は複数の手を検出できるものである.

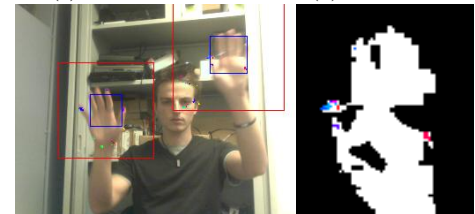
### 3.2 検出実験

実環境では、ノイズや虫食いが発生する. このような状況における本手法の有用性と課題点を調査するため指先検出実験を行った. なお、データセットの解像度は  $320 \times 240$  画素である. 検出結果を図5に示す. 指先の着色された画素は指先として検出されたものである.

実験結果より、本手法により、少々のノイズや虫食いがあっても指先位置を検出できることが分かった(図5(b))が、虫食いが激しいと誤検出が多くなるという事も分かった(図5(d)). また、指同士の間隔が狭い場合や短い指の検出が困難であることも分かった.



(a) 検出結果画像 (b) 左手詳細



(c) 検出結果画像 (d) 左手詳細

図5. 検出結果例

## 3.3 処理速度評価

実環境での任意の連続した 1,000 フレームを対象に処理速度の計測を 3 回施行した. 結果の平均処理速度を表1に示す. なお、この実験には Intel Core 2 Duo (1.07GHz), 2GB メモリの PC と BUFFALO BSW20K04H のカメラを用いた. 実験対象の解像度は QVGA( $320 \times 240$  画素), フレームレートは 30fps である.

表1. 処理速度の評価結果

前景抽出	手領域検出	指先検出
16.71ms	3.18ms	12.20ms
32.09ms		

## 4. まとめ

本稿では、単眼カメラでのハンドジェスチャー認識システムのための手指検出手法を提案した. 実験の結果から、本手法は、実時間処理のために十分高速な手法であり、前景抽出の精度がある程度良ければ、実環境での指先検出が可能だと言える.

しかし、手とカメラとの距離が近い場合は手領域候補が多くなるため時間が掛かり、手がカメラから離れ過ぎる場合にはノイズや虫食いが増えるに伴って指先の誤検出が増加するという課題がある.

また、本稿では指先検出までの提案を行ったが、後段のジェスチャー判定方法の提案・実装も今後の課題である.

### 参考文献

- [1]根本, 他, “複雑背景下における単眼カメラによるハンドジェスチャーの認識”, FIT, H-010, (2013).
- [2]佐藤, 他, “Radial Reach Filter (RRF) によるロバストな物体検出”, 信学論(D-II), vol.J86-D-II, no.5, pp.616-624 (2003).
- [3]Sato Y, et.al, "Fast tracking of hands and fingertips in infrared images for augmented desk interface", AFGR, pp. 462-467, (2000).
- [4]IPA 「教育用画像素材集サイト」  
<http://www2.edu.ipa.go.jp/gz/>
- [5] V Spruyt, et.al, "Real-Time Hand Tracking By Invariant Hough Forest Detection", ICIP (2012).