

スマートフォンによるVRのためのジェスチャ入力の実装・評価

コンピュータサイエンスコース 学籍番号:1311151 成見研究室 長谷川健吾

1 はじめに

近年, PlayStationVRなどのゲーム機器の登場により, Virtual Reality(以下 VR)の体験が身近なものになってきている。近年はPCとの接続をワイヤレス化した製品や, PCが不要のシステムなどが注目されている。

VRの体験にはHead Mounted Display(以下 HMD)が必要であり, 体験に操作が伴う場合はコントローラなどの入力デバイスを用いるのが一般的である。しかし, VRの体験にHMDとコントローラの両方を揃えるにはコストが大幅にかかってしまう。

これらの課題を加味し, 本研究ではPCが不要でスマートフォンだけで手軽に操作を伴うVR体験の実現を目指す。

2 既存製品・関連研究

2.1 Leap Motion Mobile Platform

Leap Motion社が発表した製品で, 手のジェスチャを認識しコンピュータへ入力を行うデバイスであるLeap Motionをスマートフォン向けに改良したハードウェアである[1]。この製品はスマートフォンを装着して使用するHMDであるGearVRの, 背面に搭載されたスマートフォンからの給電のみで動作し, ジェスチャの高精度な認識と, 高速な読み込み, および高画角な認識ができるという点に優れている。本研究では, 追加のデバイスを用いない点が異なっている。

2.2 HMDを用いた仮想ガジェットの開発

岡田浩臣らの研究[2]は, HMDの背面カメラからジェスチャとARマーカを認識し, 実世界の拡張を図っている。HMDの内臓カメラからジェスチャとARマーカを取り込み, PC処理による比較的精度の高いハンドトラッキングを実現している。また, HMDにWrap920ARを用いており, ワイヤレスであり内臓カメラを利用できる。本研究では, ARマーカが不要な点, 別途PCが不要な点が異なっている。

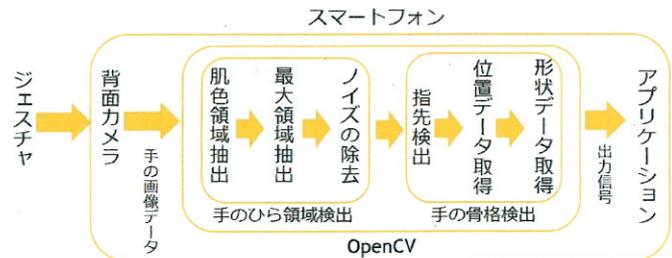


図1: システム構成

3 開発したシステム

本研究で開発したシステムの構成を図1に示す。スマートフォン中で行う操作は, 手のひら領域検出と手の骨格検出によって構成される。

手のひら領域検出では, 内臓カメラによって取得される画像をOpenCV¹を用いて手のひら領域に対しての切り出しを行い, ノイズを除去したのち, 輪郭データを出力する。

手の骨格検出では指先を検出し, 座標変換を行う。その座標を参考に形状を計算し, 形状に応じた特定の数値を出力する。

3.1 肌色領域抽出

既存手法として, 服の袖と手のひらの色の差から肌色領域を抽出する手法[3]などがあるが, 画角の問題から本研究では色相・彩度・明度からなる, HSV色空間による2値化により, 肌色領域の抽出を行った。閾値を以下に示す。

- 色相 (H) : 0~25
- 彩度 (S) : 50~255
- 明度 (V) : 50~255

この値に当てはまるHSV値を肌色領域, つまり手の領域を含む領域とした。この値は比較的誤検出が多いが, 手のひら領域が検出できる確率が高いため用いた。

¹インテルが開発・公開したオープンソースのライブラリで, 画像処理・画像解析および機械学習を行う。

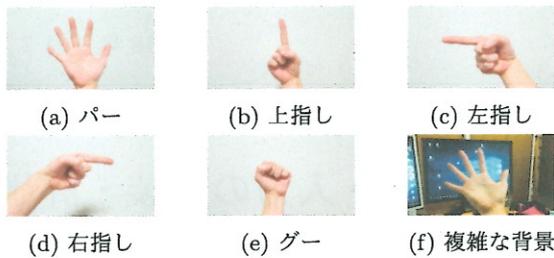


図 2: 評価実験で用いた画像

3.2 最大領域抽出

肌色領域抽出によって抽出される肌色領域は複数存在する。その中で手のひらの領域を含む肌色領域のみを抽出した。背面カメラに手のひらを映したとき、画面中最大の肌色領域は画角の関係上手の領域であると考えられる。そこで、本研究では抽出された肌色領域の面積を計算し、最大の面積を持つ領域を手のひら領域とした。

3.3 ノイズの除去

ここまでの操作では、まだ多くのノイズが残る。そこで本研究では Teh-Chin チェーンの近似アルゴリズムと Douglas-Peucker アルゴリズムの利用を考えた。前者は細かな角の部分の頂点をスキップし、より滑らかな図に変換するアルゴリズムで、後者は 2 頂点間の距離が輪郭に対する特定の割合に達しない頂点を消去するアルゴリズムである。本研究では試行の後、前者の Teh-Chin チェーンの近似アルゴリズムのみ適用した。

3.4 手の骨格検出

指先点は輪郭のなかで凸点と判定できるため、周囲の頂点との内積を比較し、凸となる頂点を指先点候補とした。候補点はノイズにより発生することもあり、ノイズで発生した際は毎フレーム大きくずれることが確認されている。そこで位置データを取得する際は、各候補点がフレーム毎に動いているかどうかで絞込を行った。これにより得られる信頼度の高い点を指先点とし、各指先点と手の重心点から、 x, y 座標に平行な直線と内積を用いて骨格を計算し、ジェスチャの推定を行った。

4 評価

Galaxy S6 edge を用いて、背面カメラで 1280×768 ピクセルの画像を取得して評価実験を行った。再現性を良くするため、パー・上と左右を指す・グーのジェスチャの 5 枚の画像 (図 2a~2e) と、背景が複雑な環境下でのパーのジェスチャの画像 (図 2f) をディスプレイに映し、それをカメラで撮影した。

表 1: HSV の値による検出誤差

範囲	色相 (H)	彩度 (S)	明度 (V)	(a)~(e) の検出面積 (%)	(f) の検出面積 (%)
本研究	0~25	50~255	50~255	97.14	52.95
狭い	0~25	60~180	80~230	60.03	19.78
広い	0~25	25~255	25~255	100.23	90.76

表 2: 近似オプションによるフレームレートの変化

近似オプション	(a) の頂点数 (個)	フレームレート (fps)
なし	4500	7.47
SIMPLE	2200	9.01
Teh-Chin チェーン	1200	14.44

4.1 HSV の値による検出誤差

色相 (H) の値を固定し、彩度 (S)・明度 (V) を本研究から狭くした場合と広くした場合で評価を行った (表 1)。

表中の検出領域は、肌色領域の正しい面積と検出した面積の割合である。

広い範囲の方が検出面積については誤差が少ないが、実際にはノイズが多く出ていたため、本研究のパラメータに設定している。複雑な背景では検出が難しい。

4.2 Teh-Chin チェーンの近似アルゴリズムによる速度の評価

Teh-Chin チェーンアルゴリズムの有無と、メモリの使用を抑えるために一般的によく使用される SIMPLE オプションの 3 種類で、計算速度の比較を行った (表 2)。SIMPLE では頂点数が半分ほどとなり、Teh-Chin チェーンでは 3 分の 1 から 4 分の 1 ほどになった。

頂点数の低下とともに、速度の向上が見られた。指先点の検出精度の低下は確認されなかった。

5 まとめ

本研究では VR アプリケーションを内臓カメラから操作することを想定した、スマートフォンによる手のひらを認識するシステムの開発を行った。

なめらかな図形である手のひらは Teh-Chin チェーンの近似アルゴリズムにより効果的に近似され、比較的誤検出を抑えてかつ速度を向上できた。

一方で、動きのあるジェスチャの認識は難しく、光による検出率の低下なども現状問題として残っている。

参考文献

- [1] "LEAP MOTION BLOG" <http://blog.leapmotion.com/mobile-platform/> (最終アクセス 2018 年 1 月)
- [2] 岡田浩臣, "HMD を用いた仮想ガジェットの開発", 高知工科大電子・光システム工学科 2012 年卒業研究報告
- [3] 玉城絵美・星野聖, "安定した肌色抽出のための個人別表色系", TVRSJ, vol.12, No.4, pp.471-478, 2007