

天井から下方に向けて設置した RGB-D センサによる 指差しポインティングの研究

野田雄希[†] 水谷晃三[†]

帝京大学理工学部情報電子工学科[†]

1. はじめに

人が行う指差しジェスチャをシステムが認識してその方向を推定することで、スクリーン上にポインタを表示する指差しポインティングの研究が行われている。環境内に設置されたセンサを用いて指差しジェスチャを認識することで、利用者がデバイスを直接使用することなくポインティングできる環境が実現される。本稿では、複数台の RGB-D センサを天井から下方に向けて設置することで、室内全体で複数人が同時使用できる指差しポインティングシステムの実現方法について述べる。

2. 研究背景

指差しポインティングに関する既存研究として、ユーザの正面に横並びに設置した2台のカメラを用いる手法[1]やユーザの上側と左側に直交するように設置したカメラ2台を用いる手法[2]などがある。いずれもカメラで捉えたジェスチャを機械学習の仕組みで認識して指差し方向を推定する。これらの方式では、複数人のジェスチャを同時に認識しようとする場合、あるユーザのジェスチャが他のユーザに遮られるオクルージョンが生じる。一方、使用位置の制限なくポインティングやスライド操作が可能なテーブルトップ型システム[3]も提案されている。しかしながら、本システムではユーザごとに機材が必要であり多人数で使用には向いていない。

これらの問題に対し、本研究では複数台の RGB-D センサを天井に下方へ向けて設置する。天井から見下ろす形でユーザのジェスチャを認識するためオクルージョンの発生を軽減できる。また、室内全体を捉えるように複数のセンサを並べて配置することで、あらゆる場所でユーザの指差しジェスチャを認識できるようにする。こうすることで、広い会議室や教室などで多人数が同時に使用できるポインティングシステムが実現できるようになると考えられる。

A Study on Hand Pointing Detection using RGB-D Sensors Placed on the Ceiling with Downward Direction

[†]Yuki Noda, Kozo Mizutani, Department of Information and Electronic Engineering, Faculty of Science and Engineering, Teikyo University

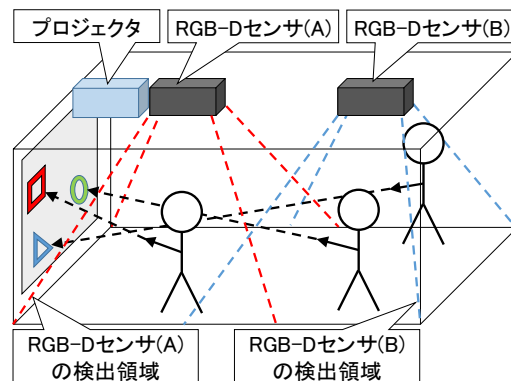


図1. システム構成図

3. 方法

3.1. システム概要

図1のように本方式では RGB-D センサを天井から下方に向けて複数台設置する。センサから取得した深度データを用いて指差しジェスチャを認識する。この際、指先と指の付け根の位置情報を取得して指が差しているスクリーン上の座標を推定し、ユーザごとにポインタを表示する。なお、指差しは人差し指のみで行い、ポインタの移動は人差し指のみまたは腕全体を動かすこととする。

3.2. ポインティング方式

指差しジェスチャにより対象物を指し示すとき、その対象物は指が差す直線の延長線上にあると仮定される。このとき、ユーザと対象物が近いほどに指差し方向と対象物の距離の誤差は小さく、距離が遠くなるほど大きくなる。本研究では室内のあらゆる場所で使用できることを想定している。スクリーンに近い位置にいるユーザと遠い位置にいるユーザでは、投影されるポインタ操作の容易性が異なると考えられる。そこで本研究では、指差しジェスチャに対するスクリーン上のポインタの動作について、直接方式と移動方向指定方式の2通りの方式を検討する。直接方式はスクリーンに向かって指差す方向の実際のスクリーン上の座標にポインタを表示する方法であり、ユーザとスクリーンの距離が近い場合を想定している。移動方向指定方式はスクリーンに対する指の角度に応じてポイン

タの移動方向を指定する方式である。指差す方向とその角度に応じてカーソルがスクリーン上を移動する。本方式はユーザとスクリーンの距離が遠い場合を想定している。

3.3. カスケード分類器による手の検出

指差しジェスチャを認識するため、本研究では OpenCV のカスケード分類器によって指差し状態の手とその位置を検出する。RGB-D センサで得られた深度データを、あらかじめ決めた深度値の範囲で画像化する。これを深度画像と呼ぶ。深度画像を分類器で処理することで画像中の手の領域 R_{hand} の位置と大きさを得る (図 2)。

カスケード分類器は画像中のオブジェクトを検出するための機械学習手法の一つである。分類器の作成にあたっては正例画像と負例画像の 2 種類の画像を事前に用意して学習を行う。本研究においては、指差しジェスチャを行っている手が含まれる画像を正例画像、それ以外の状態の画像を負例画像として用意する。

3.4. 指差し方向の推定方法

指差しジェスチャの際の方向を推定するために、カスケード分類器で検出した手の領域について、図 2 に示すようにピクセルを走査してラベリングしていき該当ユーザの人体部分のセグメンテーションを行う。人体部分のピクセルのうち、最も深度値が小さいピクセルを頭部位置 P_{head} とする。次に R_{hand} の領域中 P_{head} から最も遠い点 P_{tip} を指の先端として決定する。さらに、 P_{tip} から同心円状に半径を変えながら走査していき、人体部分のピクセル数が閾値以上になったときのピクセルの重心を指の付け根の点 P_{base} として決定する。得られた P_{tip} と P_{base} の座標および深度値を用いて指差し方向を決定する。

4. システムの試作および考察

前述した方式によるシステムの試作を行った。図 3 にシステムの実行例を示す。RGB-D センサには Kinect v2 を使用した。指差す方向にポインタが表示されることを確認できた。手の検出精度が十分でなく、検出されないときにポインタが消滅する問題が確認された。学習に用いる画像の見直しなどを行う必要がある。また、本方式ではスクリーン上のポインタの垂直方向の座標はセンサから得た深度値により決まる。深度値のノイズは平滑化フィルタにより除去しており、実際のジェスチャに対してポインタが遅れて反応する。実用的な指差しポインティングの実現のためにはこの改善も必要である。

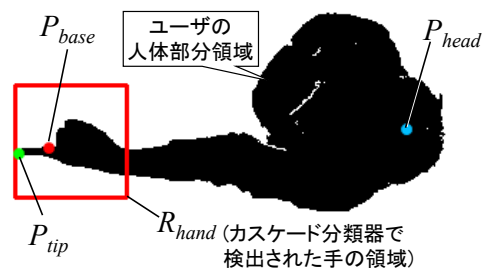


図 2. 指差し方向の推定

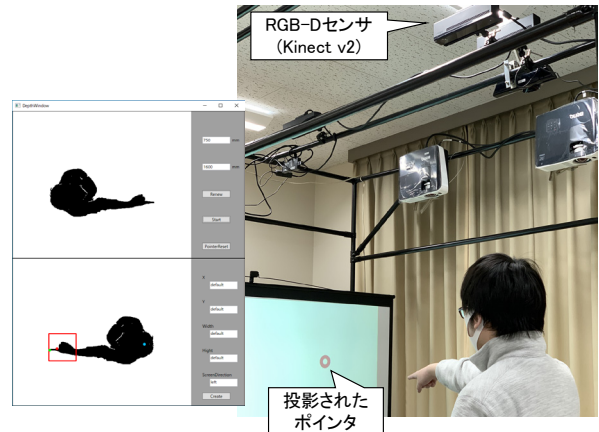


図 3. 試作システムの実行例
(左：システム画面，右：システムの設置状況およびポインタの投影例)

5. おわりに

本稿では、RGB-D センサを天井から下方に向けて設置することで、部屋全体で複数人が同時に使用できる指差しポインティングシステムの実現方法について述べた。ユーザとスクリーンの距離を考慮し 2 パターンのポインティング方法を検討し、試作システムを実装した。

- [1] Dai Fujita, Takashi Komuro: Real-time 3D Hand Pointing Recognition using Appearance Difference between Two Camera Images, The 3rd IAPR Asian Conference on Pattern Recognition (ACPR 2015) Program Booklet, pp. 36-37, 2015.
- [2] K. Hu, S. Canavan, L. Yin: Hand Pointing Estimation for Human Computer Interaction Based on Two Orthogonal Views, Proceedings of International Conference on Pattern Recognition, pp. 3760-3763, 2010.
- [3] Shun Sekiguchi, Takashi Komuro: A Tabletop Projector-camera System for Remote and Nearby Pointing Operation, Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp. 1621-1626, 2015.