

## 深度センサ付きカメラを用いた手持ち物体の3次元トラッキング 3D Tracking of In-Hand Objects using Depth-Sensing Camera

金山 知俊<sup>2)</sup>

Chishun Kanayama

### 1 はじめに

モーションキャプチャーなど、人の動作を計測し、デジタルデータとしてコンピュータに取り込み再現する手法は映画やゲーム、VR キャラクターなど様々な用途で広く行われている。また、ビデオカメラや超音波センサなどを用いて手の形や動きをジェスチャーとして取り込み、コンピュータを操作する手法も存在する。人は様々な道具を手にとって使用することで日常生活を営んでいるが、その動作をコンピュータに取り込むことを考える場合、手に持った道具を認識し、その位置姿勢を計測、追跡(トラッキング)する必要がある。

我々はコンピュータを用いて書道の動作を取り込み再現する仮想書道システムの研究を行っており [1, 2], 筆に取り付けたセンサと深度センサ付きカメラを使用することで実際の毛筆を筆記デバイスとする手法を提案している [3]. この手法では深度センサ付きカメラ(以下、深度カメラ)で筆の位置を計測し、筆に取り付けた9軸 IMU で姿勢を求めており、仮想書道の筆記に適用可能であることを確認している。ただし、この手法には取り付けたセンサが筆記動作の妨げになることや事前のキャリブレーションが必要で手軽に利用できないなどの問題がある。

そこで本研究では深度カメラのみを用いて手に持った物体の位置姿勢を計測し、その動きを追跡する手法の実現を目指しており、今回は毛筆の筆記動作を取得する手法を考案した。筆を持つ手を深度カメラで撮影し、得られた点群のうち手と筆の範囲を抽出し、RANSAC による直線推定を行なって筆の位置姿勢を求める。この処理を継続することで手に持った筆の3次元空間におけるトラッキングを実現する。

### 2 手に持った毛筆の位置姿勢計測

本研究では以下の手順で手に持った毛筆の位置姿勢を求める。

- 深度カメラによる RGB-D 画像の撮影
- 撮影画像内の手の位置計測
- 手の位置を中心とした点群抽出
- 直線推定による毛筆の位置姿勢計測

#### 2.1 深度カメラによる RGB-D 画像の撮影

本手法は机上の紙面に毛筆で筆記を行う様子の取得を対象としており、利用者を正面から撮影する位置に深度カメラを配置する。使用した深度カメラは Intel RealSense D415 である。この深度カメラの計測可能範囲はカメラ正面から 30cm 以上で、視野角は水平 69.4 度、垂直 42.5 度である。この計測可能な範囲内で筆記動作を行う必要がある。

#### 2.2 撮影画像内の手の位置計測

筆記中の毛筆は手で保持されているので、筆の一部は手で遮蔽されることになる。ただし、毛筆の軸は手の上

下に見えているので、撮影画像から手を検出し、手を中心とした領域を抽出することで毛筆のおおよその位置を求めることができる。

本研究では手の位置を求めるために MediaPipe Hand[4] を使用した。MediaPipe Hand は画像内から手を検出し、関節位置を示す 21 個のランドマークから手の姿勢を求めることができる。今回は画像から得られたランドマークの重心を手の中心位置とみなしている(図 1)。

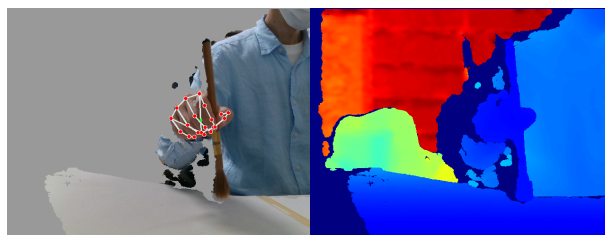


図 1 撮影画像内の手位置検出

#### 2.3 手の位置を中心とした点群抽出

深度カメラで撮影した RGB-D 画像を点群に変換し、前節で求めた手の中心位置の周囲を切り取ることで、毛筆を持った手周辺のみ点群データを抽出する(図 2)。毛筆の長さはおよそ 20cm 前後なので、手を中心とした 1 辺 25cm のバウンディングボックスを用いて抽出を行う。



図 2 撮影画像から変換された点群

#### 2.4 直線推定による毛筆の位置姿勢計測

抽出した点群には手に持った毛筆が含まれている(図 3)。毛筆は細長い円筒形をしているので、筆の軸表面の点群に近い直線の式を求めることで毛筆の位置姿勢を決定できる。本研究では RANSAC の手法で直線の式を求めている。

### 3 実験結果と考察

本研究の手法の効果を確認するために python 言語を用いてテストプログラムを作成し、実験を行なった。点

2) 南山大学理工学部ソフトウェア工学科

群データの処理には Open3D[5], RANSAC による直線推定には pyRANSAC-3D[6] を用いた。

使用した PC は Sony VAIO S, 撮影画像は VGA サイズとして RGB-D 画像を撮影したところ, 直線推定処理を除いて約 7FPS で動作した。処理性能を向上させるには点群のサブサンプリングやより高速な PC の使用が必要である。

RANSAC による直線推定を実施した結果, 毛筆を示す点群にほぼ一致する方向の直線を得ることができた(図 4)。ただし, RANSAC の処理には 1 フレームあたり約 0.5 秒の処理時間がかかっており, 現時点ではリアルタイムの処理を実現できていない。これについても点群の削減やより高速な PC の使用を検討している。

また, 直線推定により筆の姿勢とおよその位置は得られるが, 毛筆筆記に必要な筆の穂先の位置は別途求める必要がある。深度カメラで得られた点群の座標から机上の平面を求め, 穂先の点群との距離を求めることで穂先の位置を決定する予定である。

なお, 画像内の手の検出において, 指の方向の誤認識や手自体の認識失敗が発生することがある。現時点では MediaPipe Hand を用いて手の位置を求めているが, 手の姿勢自体は必要としていないので, 特徴点の抽出や追跡による手法を併用することやカルマンフィルタ等の時系列フィルタを使用することも検討している。

#### 4 おわりに

本研究は深度カメラのみを用いて手に持った物体の位置姿勢を計測し, トラッキングを可能とする手法の実現を目指しており, 今回は毛筆の筆記動作を対象とし, その位置姿勢を取得する手法を考案した。筆記動作を深度カメラで撮影して得られた RGB-D 画像を点群に変換し, 手の位置を中心とする点群を抽出して RANSAC による直線推定を行うことで毛筆の位置姿勢を求めることを可能とした。

現時点ではリアルタイム処理に十分な速度が得られていないので, 点群の削減や, より高速な PC の使用を検討している。また, 手の検出や直線推定についてもより高速な手法や頑健な手法の使用を検討している。今後, これらの課題に取り組んでリアルタイムでの処理を実現し, 筆にセンサを取り付ける手法 [3] と比較して本手法の有効性を確認する予定である。

#### 謝辞

本研究はJSPS 科研費 JS19K12290 の助成を受けたものです。

#### 参考文献

- [1] 金山知俊: 仮想書道のための柔軟な穂先をもつ毛筆モデルの提案, *FIT 2018*, (2018/9/20).
- [2] 金山知俊: 墨量の減少によるかすれ表現を可能とする仮想書道のための筆跡生成手法, *FIT 2020*, (2020/9/1).
- [3] 金山知俊, “9 軸 IMU と深度センサを用いた手書き入力デバイスの開発”, *FIT 2021*, (2021/8/27).
- [4] Mediapipe Hand, [https://developers.google.com/mediapipe/solutions/vision/hand\\_landmarker](https://developers.google.com/mediapipe/solutions/vision/hand_landmarker) (2023/6/12 参照).
- [5] Open3D A Modern Library for 3D Data Processing, <http://www.open3d.org/> (2023/6/12 参照).
- [6] pyRANSAC-3D, <https://github.com/leomariga/pyRANSAC-3D> (2023/6/12 参照).



図 3 手を中心とした点群抽出

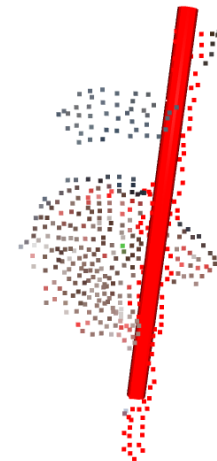


図 4 直線推定結果