

仮想世界に置けるアイテムの透過的な利用のための3次元物体のリアルタイムトラッキング

柏木 敏朗^{†1,a)} 角 薫^{†1,b)}

概要: 今日、モーションセンサーデバイスの登場により、人間の体や手の動作を容易に認識することができるようになった。しかしそれらは、事前に人間のスケルトンデータや動作を定義しているため、トラッキング対象が一意的であり人間以外の物体の動作を認識することができない。そこで本研究は、事前に対象の物体の情報を定義せずに、リアルタイムに対象の物体のスケルトンデータを取得する。そのために対象の物体の関節の位置を認識する。まずセンサーカメラを用いて距離画像を取得し、コンピュータビジョンを用いて物体の形状の解析を行う。そこで解析された情報をもとにコンピュータビジョンと3次元点群処理を用いて動作点と支点を識別する。また得られたスケルトンデータを用いて対象の物体のポインタラッキングを行う。

Real-time 3D Object Tracking for Transparent use of Items in the Virtual World

TOSHIRO KASHIWAGI^{†1,a)} KAORU SUMI^{†1,b)}

Abstract: In recent years, motion sensor devices made it possible to easily recognize human body movement and hand gestures. However, since these devices define human skeleton data and motions in advance, tracking targets are limited, and it is impossible to recognize the motion of non-human objects. The purpose of our research is to acquire skeleton data of target in real time without defining the information of the object in advance. Therefore, the system recognizes the position of the joint of the target object. First, depth images are obtained with sensor camera, and the shape of target is analyzed by using computer vision technology. Based on analyzed data, operating points and fulcrum is identified by using computer vision technology and three-dimensional point cloud processing. In addition, target object is tracked by using obtained skeleton data.

1. 背景

近年、3D 深度センサーなどのセンサー技術の発達により、マーカーレスの物体の3次元センシングの研究が注目されている。それらの研究の中に物体のモーショントラッキングと呼ばれる研究がある。物体の位置や回転情報、関節の3次元座標を計測を行う。従来から存在するモーションキャプチャは、複数のカメラでマーカーが装着された物体を囲むことで計測が行われていた。しかし深度セン

サーによるモーショントラッキングにより、マーカーを装着せずに1つのセンサーデバイスで行うことができるようになった。具体的なセンサーデバイスにKinectとLeap Motionが挙げられる。それらのセンサーデバイスの前でユーザーの体や指を動かすことで、容易にポインタラッキングを行うことができる[1][2]。これにより、ジェスチャー操作などのNUI^{*1}を実現した。

2. 目的

KinectやLeap Motionなどのモーションセンサーデバイスはそれぞれに対象となる物体が定められている。例え

^{†1} 現在、はこだて未来大学
Presently with Future University Hakodate

^{a)} b1014090@fun.ac.jp

^{b)} kaorus@fun.ac.jp

^{*1} Natural User Interface

ば、Leap Motion であれば手がトラッキング対象として定められている。なぜなら、事前にトラッキング対象となる物体のスケルトンデータを学習されており、それに応じた推定システムが組み込まれているからである。そのため、モーションセンサーデバイスを用いたモーショントラッキングは人間の体や手などのトラッキング対象に限定されている。そこで本研究の目的は、物体の動作からその物体の関節を識別するシステムの開発を行うことである。本研究により、センサーデバイスによるポイントトラッキングを複数の物体を対象に行うことができるようになると思われる。

3. 関連研究

3.1 事前知識なしでのモーショントラッキング

従来研究に事前に対象物体のスケルトンデータなどの情報を学習せずにモーショントラッキングの研究がある [3]。これはレーザーセンサーを用いて行われ、任意の物体が対象である。レーザーの照射点におけるターゲットの位置の位置計測により物体のモーショントラッキングを実現している。しかし私の研究は物体の位置や回転情報だけでなく、物体の関節の検出とそれの 3 次元座標の計測を行う。

3.2 コンピュータビジョンを用いた動作認識

センサーカメラから得られた距離画像やカラー画像をコンピュータビジョンを用いて、物体の輪郭などの情報を抽出する。また画像に対して距離変換を行うことで物体のスケルトンを求めることができる。これらの情報の変化で物体がどのように動作しているかを認識することができる [4]。具体例として、コンピュータビジョンで指を挙げている本数を計測することでハンドジェスチャーの認識が可能である [5]。本研究ではコンピュータビジョンによる解析をもとに、物体の動作に伴う形状の変化を識別する。

3.3 点群処理を用いた物体の形状認識

点群と呼ばれる点の集合を用いて、物体の形状を認識する。点群は物体の表面を点の集まりで出力することが可能である。そのため物体の形状を点群処理を用いて解析することができる。例として、物体の部分ごとのクラスタリングを行うことができる [6]。また点群による物体検出を行うこともできる。本研究では、物体の 3 次元の形状の取得や関節ごとの識別に利用する。

4. 提案手法

本章では、対象物体のスケルトンデータをリアルタイムで解析する方法を示す。ここで対象物体とは、ハサミやスプレーなどの支点が一つである物体を指す。下記に開発環境を記載する。

- 使用するデバイス

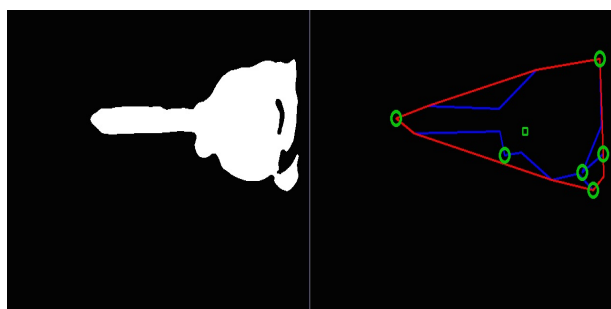
- Real Sense SR300
- 開発ライブラリ
 - Open CV
 - PCL (Point Cloud Library)
 - librealsense

対象物体のスケルトンデータを取得するために、対象物体の関節を推定する。その推定方法は 3 つの工程がある

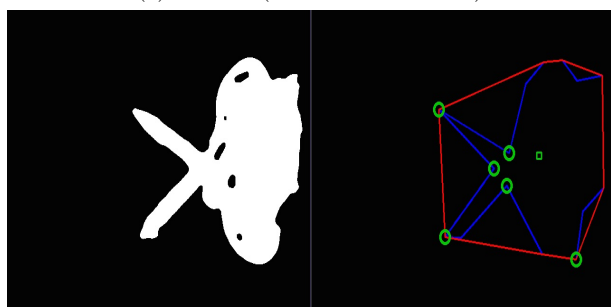
- (1) 対象物体の形状を解析
- (2) データから動作点を識別
- (3) 動作点をもとに支点を推定

4.1 対象物体の形状を解析

対象物体の形状を物体の動作から解析を行う。まず対象物体を動かし、RealSense から 2 つのポーズごとの距離画像を取得する。例として、対象物体がハサミの場合、図 1 のように開いたときと閉めたときの 2 つのポーズを行う。それぞれのポーズごとの距離画像から対象物体の輪郭を検出する。そこから対象物体の簡略な形状を認識するために得られた輪郭に直線近似を適用する。その結果から対象物体の輪郭の変曲点を求める。変曲点は簡略された輪郭の頂点の中で鋭角である頂点とする。次に、得られた輪郭に対して凸包を行い、凸集合を求める。この結果を図 1 に表す。図 1 は (a) と (b) で 2 つのポーズを表している。それぞれの図の左図は Real Sense から得られた距離画像であり、右図は物体の形状解析の結果である。右図の対象物体の輪郭の周りにある線は、凸包で得られた凸集合を結んだ線である。また円は輪郭の変曲点を表し、四角は凸集合の重心を表している。



(a) ポーズ 1(はさみが閉まる場合)



(b) ポーズ 2(はさみが開く場合)

図 1: 2 つのポーズの対象物体の形状データ

その後、その凸集合の重心を原点とする変曲点の相対座標を求める。この一連の流れを 1 秒間行う。そこで得られた変曲点の相対座標の散布図を図 2 に示す。丸がポーズ 1 の変曲線の相対座標を、三角形がポーズ 2 の変曲線の相対座標を表している。このグラフから対象物体は x の相対座標の -270 から -200 の間の変曲点が動作していると予想できる。またその x の変位で縦方向に動作する物体であると予想できる。

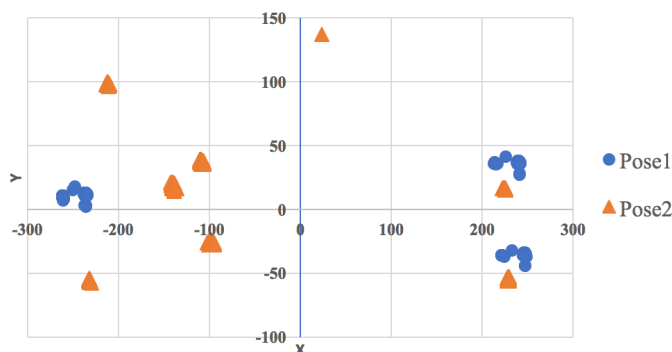
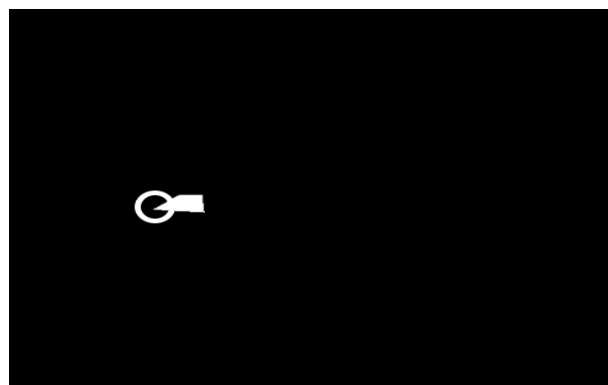


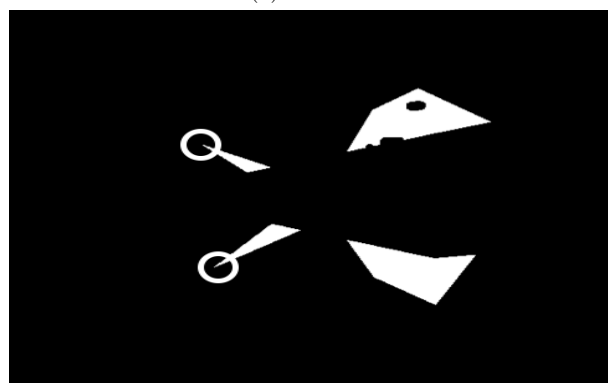
図 2: ポーズごとの相対座標

4.2 データから動作点を識別

先ほど得られた 2 つのポーズごとのデータを用いて動作点を識別する。先ほど 1 秒間に得られたポーズごとの相対座標に x-means 法 [7] を用いてクラスタリングを行い、散布図の点を部分ごとに集約する。まず、ポーズ 1 とポーズ 2 の輪郭に対して差分を行う。次に、ポーズ 1 とポーズ 2 の輪郭に対する凸集合の差分を取る。下図のように、これら 2 つの結果の共通部分を抽出する。変曲点の中で、2 つの差分の共通部分に近く、凸集合の重心よりも左側にある点を抽出する。この点が対象物の動作点となる。図 3 にその結果を示す。図 3 の白塗りの部分が 2 つの差分結果の共通部分を表しており、丸が推定された動作点である。



(a) ポーズ 1



(b) ポーズ 2

図 3: 共通の差分結果と動作点

4.3 動作点をもとに支点を推定

認識した動作点を用いて支点を識別する。まず、2 つのポーズの点群データを取得する。次に、それぞれの動作点の位置から近い点群を一定量分抽出する。その後、図 4 のように、それぞれの抽出された点群に対し最小二乗法を行い、直線化する。これは支点へ向かう直線を求めている。それらの直線が密集する部分を対象物の支点として識別する。

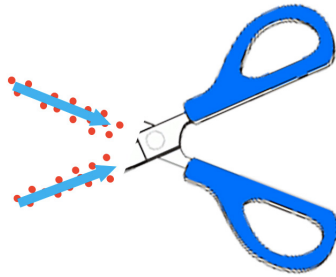


図 4: アルゴリズムのイメージ図

図 5 に関節推定の結果を示す。Real Sense で得られた距離画像を表しており、関節と判断された部分を円で囲んでいる。

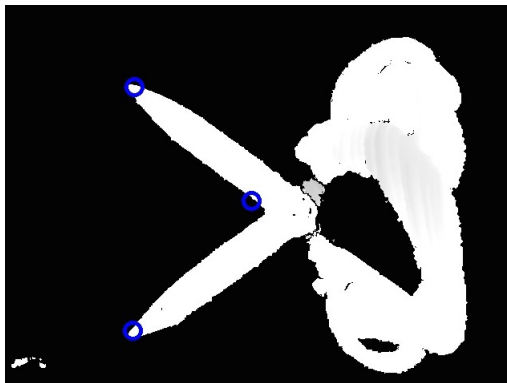


図 5: 関節推定の結果

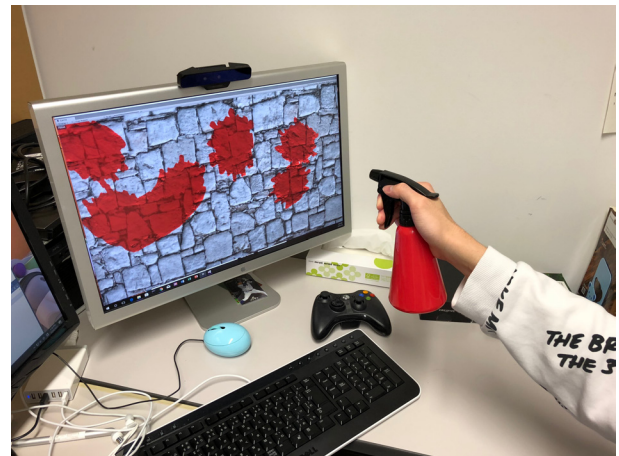
5. 応用

本研究のシステムで取得した対象物体のスケルトンデータを用いてボートラッキングを行う。つまりセンサーカメラの前で物体を動かすだけで、その物体を対象としたボートラッキングを行う。またこれをゲームなどの3次元コンテンツと連携することで、現実世界と同じように仮想世界上で現実世界の物体の操作をシミュレーションすることができる。例えば、ゲーム上のアイテムが現実世界の物体であるゲームの作成を行うことができる。ゲームが進行する中で障害となる物体や敵が現れたが、その障害に対するアイテムを所持していないとする。その時、現実世界の物体を用いることで、その障害を回避することができる。図 6(a) は森のフィールド上で進行を妨げるツタがあるため、それを現実世界のハサミで切断している様子である。またゲームだけでなくクリエイティブなツールとしても応用可能であると考え。現在、ペンタブレットのようなコンピュータ上に絵を描くことができるデバイスや VR と専用のコントローラーを用いた仮想空間上に絵を描くことのできるコンテンツが存在する。しかし本研究のシステムを用いることで現実世界の物体をコントローラーとした、仮想空間上に現実世界と同様の作業が可能なコンテンツが作成できると考える。図 6(b) はスプレーを用いて絵を描い

ている様子である。またインタラクティブシステムの開発により、多くの分野に用いることができると考える。ユーザーは現実世界の物理現象を学習することができるため、教育現場や専門技術の習得などの学習支援システムに応用できる。



(a) ハサミ



(b) スプレー

図 6: 本研究によるトラッキングシステム

6. まとめ

4 章より対象の物体の 2 つのポーズを抽出することでその物体の関節を推定することが可能になった。本研究で推定された関節の位置は、実際の関節の位置と若干違いがあるため、システムの改善をする必要がある。また、本研究では支点が 1 つである物体に制約して開発を行った。今後の課題として対象となる物体の制約を広げても適用できるようなシステムの改良が求められると考える。また 5 章でも示したとおり、本研究のシステムを用いたボートラッキングを開発することで現実世界の物体を仮想空間上で操作することが可能となる。

参考文献

- [1] W.Zeng : *Microsoft Kinect Sensor and Its Effect*, Multimedia at Work (2012)
- [2] A.G.Kirk and J.F.O' Brien, D.A.Forsyth : *Skeletal Parameter Estimation from Optical Motion Capture Data*, LComputer Vision and Pattern Recognition (2005).
- [3] Leo Miyashita, Ryota Yonezawa, Yoshihiro Watanabe, Masatoshi Ishikawa: *3D Motion Sensing of any Object without Prior Knowledge*, ACM Transactions on Graphics, Vol. 34, No. 6, Article 218 (2015).
- [4] S.Rautaray and M.Pandey: *Performance Analysis of Vision based Adaptive Hand Gesture Recognition System for Human Computer Interaction*, ACM New York (2014).
- [5] W.Shen and K.Zhao, Y.Jiang, Y.Wang, Z.Zhang, X.Bai: *Object Skeleton Extraction in Natural Images by Fusing Sale-Associated Deep Side Outputs*, Computer Vision and Pattern Recognition(CVPR), 2016 IEEE Conference on (2016).
- [6] Chun-Hao Huang and Edmond Boyer, Slobodan Ilic: *Robust Human Body Shape and Pose Tracking*, 3D Vision 3DV 2013, 2013 International Conference on (2013).
- [7] 石岡 恒憲: クラスター数を自動決定する *k-means* アルゴリズムの拡張について日本計算機統計学会, 計算機統計学 18(1), 3-13 (2006)

—