

1X-06

回転軸の指定による点群からの関節体のモーション再構成

進士 さくら 藤代 一成
慶應義塾大学 理工学部情報工学科

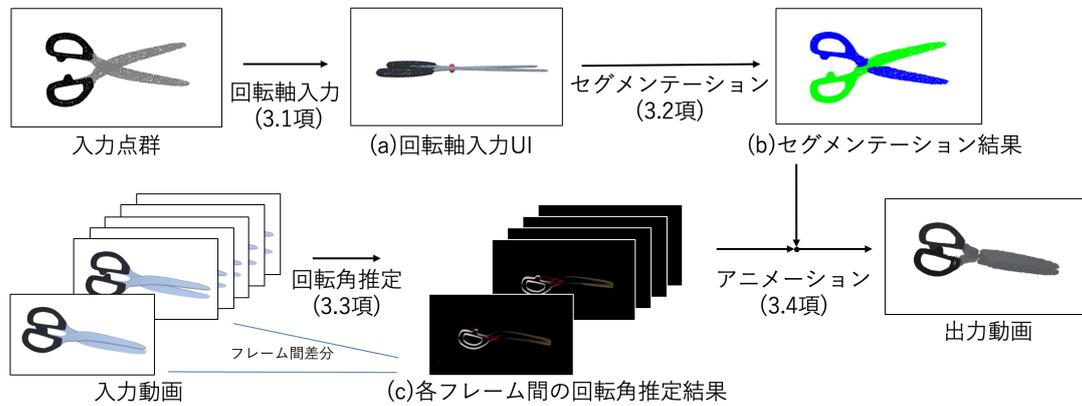


図 1: 回転軸の指定と関節運動の動画による例示をもとに、対象物の関節運動を再構成する提案手法の手順

1 背景と目的

近年、インターネット上の三次元仮想空間である、メタバースに向けて、さまざまなコンテンツが生成されている。メタバースでは、現実空間の物理的な制約を受けることなく、これまでにない新しい体験を提供できる。現実空間に存在する物体をメタバース向けに三次元再構成し、現実空間のコピーであるデジタルツインを構築すれば、観光や展示会など実存する物体に特有の内容も擬似体験できる。

しかし、三次元再構成で得られるモデルは静的であり、道具の操作などの動的な相互作用を伴う体験を実現できない。点群を連続撮影すれば動的シーンも再現できるが、撮影コストが高いうえ、各点の相関がフレーム間で維持されないため認識が困難である。三次元再構成された動的シーンを効率よく表現するには、対象の動作を理解し、得られた点群を後から適応的に変形することがより効果的である。

そこで本研究では、ある姿勢の静的な点群から、物体の運動を撮影した動画をもとに三次元の動的なシーンを再現する。可動部をもつ物体として、関節周りの運動(以下、関節運動とよぶ)を伴う関節体に注目する。

関節体の運動を推定する問題に対して、文献 [1] では、物体の関節運動前後の点群に対して差分をとることで回転情報を推定しているが、前後の状態の点群を取得するための撮影コストが大きい。文献 [2] では、点群の各パーツの形状を PointNet++ [3] によって学習し関節運動を推定しているが、未学習の物体に対応できない。

本稿では、回転軸の指定と関節運動の動画による例示をもとに、関節運動を再構成する手法を提案する。

2 概要

本研究では、一度基本姿勢の点群をスキャンした後に回転軸を入力し、使用時の様子を撮影するだけで、操作可能なモデルが得られることを目指す。本手法の手順の概要を図 1 に示す。まず上段のように、入力された回転軸を制約として点群をパーツに分割する。次に下段左のように、動画のフレーム間差分をとることで回転角を推定する。最後に各パーツを独立に回転することで関節運動を再現する。

本稿で対象とする関節体は、2つのパーツで構成され、接合部が 1 自由度の回転関節である関節体と仮定する。

3 手法

本節では、手法を図 1 に示す 4 つの手順に分割して説明する。

3.1 回転軸入力

ユーザは自由に視点を変更しながら点群に対して回転軸の始点と終点を入力する。各点を球として表示し、クリック操作で選択された球の中心座標を求めることで始点と終点を得る。図 1(a) に、回転軸を赤で強調した様子を示す。

3.2 パーツセグメンテーション

点群と入力された回転軸をもとに、点群を2つのパーツ A, B に分割する。本手法では、Random Sample Consensus (RANSAC) を用いて、回転軸を含む平面を検出することによって点群を分割した。RANSAC は、多くの誤差を含むデータに対して、任意の関数への最適な近似式を得るためのアルゴリズムで、点群に対する形状検出 [4] に用いられる。

まず、RANSAC によって回転軸を含む平面のうち最も多くの点を含む平面 Π_0 を検出し、平面 Π_0 から距離 d 内にある点群 A を抽出する。同様の手順で残りの点群に対して平面 Π_1 を検出し、距離 d 内にある点群 B を抽出する。最後に、点群 A, B のどちらにも含まれなかった点のうち、各点群からの最短距離が一定距離以下の点を分類することを繰り返し、分類済みの点を全域に拡張する。

ハサミに対してセグメンテーションした結果を図 1(b) に示す。点群を2枚の刃に分割できていることが分かる。

3.3 回転角推定

入力動画の各フレーム間において、パーツ間の角度がどのくらい変化したかを算出する。入力する関節運動の様子を撮影した動画は、回転軸上に固定したカメラから撮影され、対象の関節体以外は映っていないものと仮定した。また、パーツは片方だけが回転し、もう片方は固定されている。

まず、入力動画の各フレーム間のオプティカルフローを検出する。次に、オプティカルフローに直交する直線が最も多く通る交点 O を画像平面中の回転中心とする。最後に、交点と交点に対して直交するオプティカルフローのなす角を算出し、平均をとった角度を各フレームにおける角変化量とする。

ハサミの動作を撮影した動画の2フレーム間の差分における回転角を推定して可視化したものが図 1(c) である。

3.4 アニメーション生成

セグメンテーションされた一方のパーツを、入力された回転軸周りに、動画から推定された回転角に応じて回転させることで関節運動を再現したアニメーションを生成した。視認性を高めるため、各パーツの点群はポアソン表面再構成 [5] によりポリゴンメッシュに変換されている。

4 実装と結果

回転軸入力のユーザインタフェースは Unity で作成し、パーツセグメンテーション及び回転軸推定は Python で実装した。

ハサミと蓋付きの箱を対象に、図 1 の手順に従って関節運動を再構成した結果を図 2 に示す。入力動画と再現アニメーションの代表的なフレームの比較と、セグメンテーションによるパーツ推定結果を示した。関節運動が視覚的に尤もらしく再構成されていることがわかる。



図 2: 入力動画の特徴的なフレームと、同フレームの推定結果の比較

5 結論と今後の課題

本稿では、関節体の点群と関節運動の動画を入力とし、モデル上にユーザが回転軸を指定することによって、関節体の三次元関節運動を再現する手法を提案した。これにより、一回の点群スキャンと動画による例示で三次元関節運動が得られた。

本手法は、物体固有のパーツ形状を学習することなく回転運動が推定できるため、データが少ない民具などの文化財の使用法を体験するコンテンツ生成に有用である。本手法を多様な物体に対して適用するためには、複数の関節をもつ関節体や、複雑な形状のパーツにも対応できるセグメンテーション手法を検討する必要がある。また、本稿では、問題を簡略化するために入力動画にさまざまな制約を加えた。今後は、自然に撮影した現実の動画を用いることを目指し、回転軸上以外から撮影された画像のカメラ位置補正、操作する手や背景と対象物の分離に取り組む。

謝辞

本研究の一部は、令和4年度科研費基盤研究(A) 21H04916 の支援により実施された。

参考文献

- [1] X. Huang, I. Walker, and S. Birchfield: "Occlusion-aware reconstruction and manipulation of 3D articulated objects," in *Proceedings of the IEEE International Conference on Robotics and Automation 2012*, pp. 1365–1371, 2012.
- [2] X. Li, H. Wang, L. Yi, L. Guibas, A. L. Abbott, and S. Song: "Category-level articulated object pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2020*, pp. 3703–3712, Seattle, WA, USA, 2020.
- [3] C. R. Qi, Li Yi, Hao Su, and L. J. Guibas: "PointNet++: Deep hierarchical feature learning on point sets in a metric space," in *Proceedings of Conference on Neural Information Processing Systems (NIPS) 2017*, Long Beach, CA, USA, 2017.
- [4] R. Schnabel, R. Wahl, and R. Klein: "Efficient RANSAC for point-cloud shape detection," *Computer Graphics Forum*, Vol. 26, No. 2, pp. 214–226, 2007.
- [5] M. Kazhdan, M. Bolitho, H. Hoppe.: "Poisson surface reconstruction," in *Proceedings of the fourth Eurographics Symposium on Geometry Processing*, pp. 61–70, 2006.