



手作業における柔軟な物体表面とその接触領域の 3 次元復元

3D Reconstruction of the Object Surfaces and Contact Areas in Hand-Object Interactions

三浦康平, 岩井大輔, 佐藤宏介

Kohei MIURA, Daisuke IWAI, Kosuke SATO

大阪大学 大学院基礎工学研究科 (〒 560-8531 大阪府豊中市待兼山町 1-3,
kouhei.miura@sens.sys.es.osaka-u.ac.jp)

概要: カメラにより手作業を 3次元データとして復元することで, VR/AR 技術を用いた熟練者の手作業の追体験・学習が可能になる. 任意の対象物体に対し, 実際に手が触れていた物体表面上の接触領域を物体にマッピングすることで, 物体と手の相対位置を推定する精度が向上し, より現実に即した 3次元データが復元される可能性がある. 本発表では, 手作業における任意の物体表面とその表面上の接触領域を復元する手法を提案する.

キーワード: 3次元復元, 3次元コンピュータビジョン

1. はじめに

人の手作業は, 日常生活にはもちろん, 芸術やシェフの調理といった専門的な技術にも不可欠である. これらの手と物体との相互作用は, 手-物体間インタラクション (HOI: Hand-Object Interactions) と呼ばれる. 手の動作を阻害しない画像センサを用いて得られる 2D 画像から, HOI における手と物体を 3D 時系列データとして正確にコンピュータ上に復元することで, リアルな CG, 芸術家や料理人の造形スキルの解析と保存 [1] が可能である. さらに, 復元データの再現によりロボットハンドによる自動化 [2] や, バーチャルリアリティ技術による技術継承 [3] 等の多岐にわたる応用が可能である. そのため, HOI を正確に復元・再現することは重要な課題として研究が盛んに行われてきた [4].

HOI の中でも, 特に粘土や生地などの物体を対象とした形状変化を伴う HOI は, 人の手の特有の造形技術が発揮される. よってその形状変化の正確な復元により, 人の手作業を模倣したロボットハンドによる造形や, 伝統工芸技術などの保存と継承につながる事が期待される. このような形状変化を伴う HOI において, 形状変化を生じさせるのは

手と物体の接触であるため, 物体の変形が生じた領域の変形過程を正確に推定するには接触情報が活用できると予想される. しかし, 一般的な画像センサである RGB カメラで HOI を撮影し, その画像から復元する場合は, 物体表面が手により遮蔽されてしまうため, 手と物体の正確な接触領域が不明である. この問題には, 熱カメラの使用が有効である. HOI を熱カメラで撮影することで, 手と物体の接触による物体表面上の温度上昇を観測することができ, 温度が上昇した接触領域を取得することが可能になる. 実際, HOI の中でも, 把持姿勢の推定のために熱カメラによる接触領域の取得が活用されている [5, 6]. しかしこれらの手法は, 接触領域を剛体 3D モデルに付加し, 静的な把持姿勢を推定するため, 形状変化などの動的要素を含む HOI には適用できない.

よって, 本稿では物体形状変化を伴う HOI における, 複数台の RGB カメラと一台の熱カメラを用いた対象物体表面とその接触領域を復元する手法の提案と検証を行う. 複数台の RGB カメラにより対象物体表面を, 1 台の熱カメラにより接触領域を復元する.

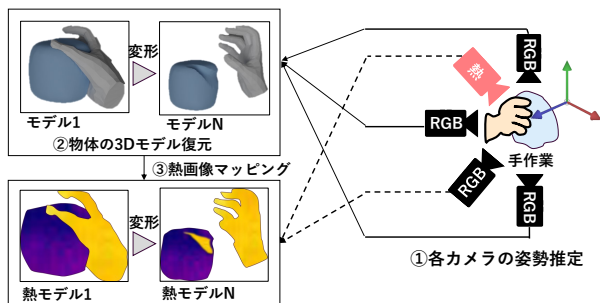


図 1: 提案手法の概要

2. 提案手法

複数台の RGB カメラと 1 台の熱カメラを用いて, 手作業における対象物体表面とその接触領域を復元する手法を示す (図 1). まず, カメラキャリブレーションにより同一の 3次元空間内におけるすべてのカメラ姿勢を推定する. 次に, RGB カメラの姿勢をもとに, HOI における物体の 3D モデルを復元する. そして, 復元した 3D モデルに熱画像を投影テクスチャマッピングすることで接触領域の情報を 3D モデルに付加する. 3D モデルの復元と熱画像のマッピングをフレームごとに繰り返すことで, 接触領域の情報を持った時系列 3D モデルを得る.



図 2: 撮影デバイス

2.1 カメラ姿勢推定

カメラ姿勢の推定のためには、まず、各種カメラのキャリブレーションにより、RGB カメラおよび熱カメラの内部パラメータを推定する。次に、その内部パラメータを用いて、既知のパターンを埋め込んだ校正儀を基準とする外部パラメータ、すなわちカメラの姿勢を推定する。なお、以降の処理を校正儀を基準とするすべてのカメラで共通の世界座標で行うために、RGB カメラと熱カメラいずれからも観測できるパターンを設計する必要がある。

2.2 対象物体表面の復元

2.1 節で推定した複数の RGB カメラの姿勢をもとに、ステレオ視で得られる深度マップを統合することで手と物体の 3D モデルをメッシュとして復元する。

2.3 接触領域の復元

2.2 節で復元した物体表面上に、2.1 で推定した熱カメラのパラメータを用いて熱画像を投影テクスチャマッピングすることで、物体表面上に接触領域を復元する。投影テクスチャマッピングとは、プロジェクタから光を投影するようにしてテクスチャを 3D モデルにマッピングする手法のことである。カメラとプロジェクタは光の出入りが逆なだけであり、幾何的なモデルはほぼ同一であるため、今回は熱カメラをプロジェクタとみなすことで実現する。

一般的に、熱画像は 1 ピクセル当たり 16 bit × 1 チャンネルの情報をもつのみでありグレイスケールである。よって、可視化のために視覚的に均一な (perceptually uniform) カラーマップを適用し色をつける。

3. 実装

RGB カメラには GoPro Hero11 Black を 14 台、熱カメラには Optris PI450i (O80 レンズ) を 1 台用いて、アルミフレームに姿勢を固定し撮影デバイスとした。各カメラは 3 行 5 列の形状で配置され、2 行目の 3 列の中央カメラが熱カメラである。GoPro は 30 fps、解像度 5312 × 2988、レンズはリニア、HyperSmooth はオフにして撮影し、熱カメラは 27 fps、解像度 382 × 288 で撮影した。

3.1 カメラ姿勢推定

校正儀として、黒塗りのアルミ円盤をベニヤ板の既知の位置に埋め込んだサークルグリッドパターン (図 3) を用

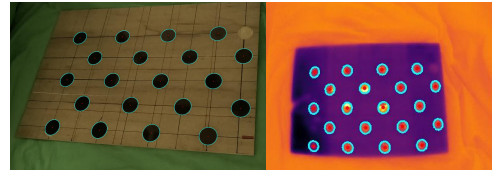


図 3: 校正儀とパターンの検出結果 (左: RGB カメラ, 右: 熱カメラ)

いた。アルミニウムなどの金属は一般にベニヤ板より赤外放射率が低いため放射熱が生じにくく、熱カメラでもその放射率の差によるアルミ円盤とベニヤ板領域のコントラストが観測できる。そのため、いずれのカメラからでもサークルグリッドパターンが検出可能である (図 3)。

カメラの内部パラメータの推定は、上記の校正儀を複数の姿勢で撮影し、それぞれの姿勢で撮影した画像における各円の 2 次元中心座標と、その 3 次元空間における座標の対応を OpenCV の `calibrateCamera` 関数に入力することで実現した。RGB カメラ、熱カメラいずれも 25 枚分の画像を使用した。なお、RGB カメラは複数台存在するが同一製品であるため、内部パラメータは一致しているとする。

次に、アルミフレームに各カメラを固定し、固定されたカメラの外部パラメータを推定する。これは 3 次元空間における基準座標系からの相対位置と回転を表すため、同一製品の RGB カメラであっても異なる。外部パラメータは校正儀を固定してこれを基準座標系として撮影し、それぞれのカメラで得られた画像について、各円の 2 次元中心座標と、その 3 次元空間における座標の対応を OpenCV の `solvePnP` 関数に入力することで実現した。

3.2 対象物体表面の復元

3.1 節で推定した複数の RGB カメラの姿勢を Agisoft 社の Metashape にインポートし、その姿勢をもとに Surface Reconstruction 処理 [7] を実行することで、3D モデルを得る。校正儀を撮影したシーンを復元した結果を図 4 左に示す。

3.3 接触領域の復元

今回、投影テクスチャマッピングを実現するには、3D モデリングソフト Blender の Projector アドオンを用いた。このアドオンでは画像をレンズ歪みがないものとして直接マッピングするため、熱画像の歪みを先に補正しておく必要がある。3.1 節の `calibrateCamera` 関数から得られる内部パラメータとレンズ歪み係数を OpenCV の `undistort` 関数に与えることで、画像の歪みを補正した。Projector アドオンの `throw` の値は 3.1 節で推定した熱カメラの内部パラメータのうち X 軸方向の焦点距離 (265) を熱画像の幅 (382) で割ったものを用いた。3.2 節で復元した校正儀の 3D モデルに、歪みを補正した熱画像をマッピングした結果を図 4 中央に示す。若干左にずれていることがわかるため、手動で仮想プロジェクタを Y 軸方向に 1 度だけ動かした。

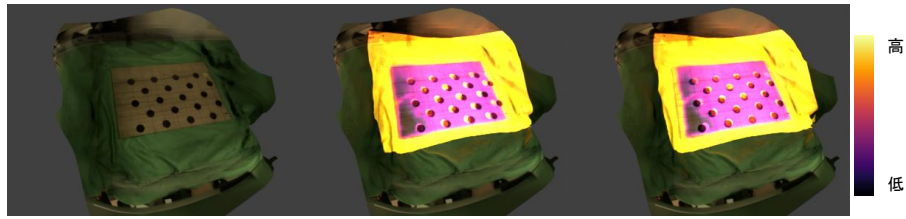


図 4: 提案手法により復元した校正儀の 3D シーン (左: RGB のみ, 中央: 熱画像マッピング, 右: 手動補正後)

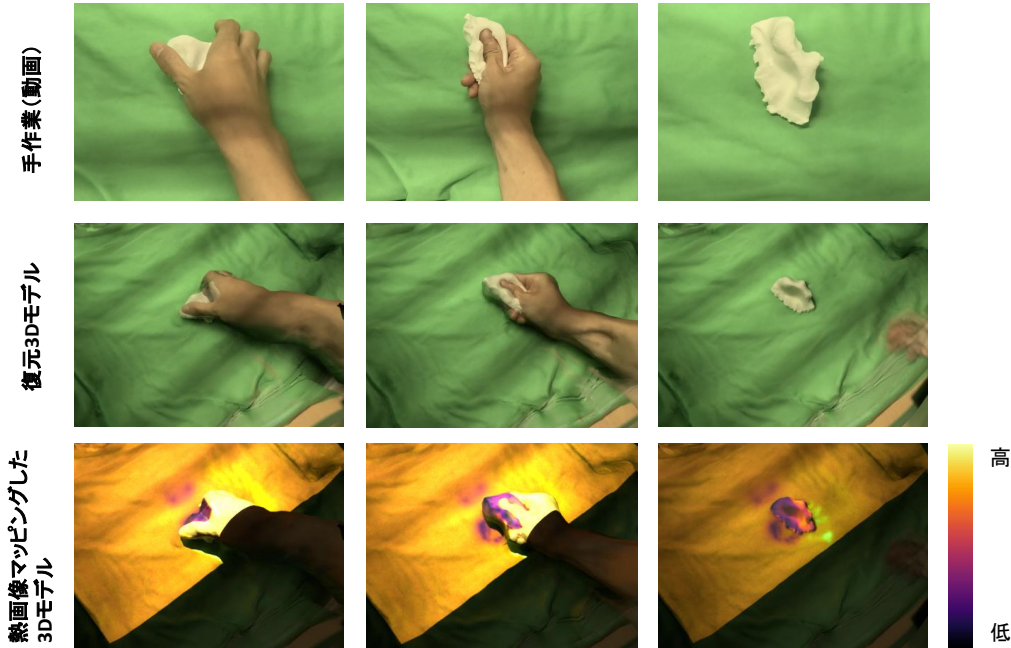


図 5: 粘土を片手でこねる手作業の復元結果. 左から右へ時系列が進む (上段: 元とした動画データの 1 カメラ視点の数フレーム, 中段: 復元した RGB テクスチャを付加した 3D モデル, 下段: 熱画像のマッピング結果)

4. 実験

実験として、粘土を片手でこねる手作業 (図 5 上段) を復元した。復元した 3D モデルは同図中段に示す。原理検証のために、図 2 にある通り上部にのみカメラを配置しているため、腕と床面の間など (図 5 の中段左赤枠部) の領域を観測できるカメラの数は少ない。そのため、そのような上部に設置したカメラから観測しにくい、あるいはできない領域は 3 次元情報が少なく、3D モデルの復元精度が低下し形状に歪みが生じている。

また、熱画像マッピングの結果を図 5 下段に示す。熱画像マッピング結果における手の形が、中段の RGB テクスチャにおける手の形と一致していることが確認できる。最右の画像においても、握られて変形した箇所が黄色く可視化されていることが確認できる。

5. おわりに

本稿では、形状変化を伴う HOI のより正確な復元のため、物体表面とその表面上の接触領域を復元する手法を提案し、紙粘土の変形シーンを撮影し検証した。仮想プロジェクトの姿勢推定精度が甘く、熱画像マッピングの結果が芳しくなかったため、今後 Blender 上のシンプルなプロジェクトによるシミュレーションではなく、3DCG ライブラリを用い

た熱画像のエッジと形状のエッジに合わせるようカメラ位置を最適化する処理などを実装することで、熱画像マッピングの精度を向上させる。そのうち、熱画像マッピングの精度を定量的に確かめるために、絵の具などを塗った手で触れた結果と、熱画像マッピングの結果間で接触領域を比較する。

参考文献

- [1] de Charette, Raoul, and Sotiris Manitsaris. "3d reconstruction of deformable revolving object under heavy hand interaction." arXiv preprint arXiv:1908.01523 (2019).
- [2] Lakshminpathy, Arjun, et al. "Contact transfer: A direct, user-driven method for human to robot transfer of grasps and manipulations." 2022 International Conference on Robotics and Automation (ICRA). IEEE, 2022.
- [3] Wu, Min-Yu, et al. "Hand pose estimation in object-interaction based on deep learning for virtual reality applications." Journal of Visual Communication and Image Representation 70 (2020): 102802.
- [4] Fan, Zicong, et al. "ARCTIC: A dataset for dexterous bimanual hand-object manipulation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.
- [5] Brahmabhatt, Samarth, et al. "ContactPose: A dataset of grasps with object contact and hand pose." Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIII 16. Springer International Publishing, 2020.
- [6] Grady, Patrick, et al. "Contactopt: Optimizing contact to improve grasps." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.
- [7] Poliarnyi, Nikolai. "Out-of-core surface reconstruction via global TGV minimization." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.
- [8] Projector Add-on for Blender <https://github.com/0cupe/Projectors>