



# 動的プロジェクションマッピングのための 輪郭に基づいた位置姿勢推定手法の精度評価

森久保優輝, 橋本直己

電気通信大学 大学院情報理工学研究科

(〒182-8585 東京都調布市調布ヶ丘 1-5-1 西 9-603, {morikubo, naoki}@hashimoto.lab.uec.ac.jp)

**概要**: 近年注目を集めている動的プロジェクションマッピングでは, 対象物体の位置姿勢をリアルタイムで推定する必要があり, 筆者らは物体輪郭を用いる手法の有効性を示してきた. さらに, この手法を基に輪郭の連続性を利用し, 背景や操作者に対して頑健化を図る手法を提案している. しかし, その精度や頑健性について十分な評価が行われていない. 本稿ではモーションキャプチャを ground truth とした定量的な評価を行い, この手法の有効性を検証する.

**キーワード**: 動的プロジェクションマッピング, エッジベーストラッキング, 精度評価, 頑健性

## 1. はじめに

近年, 自由に動かすことのできる実物体に対して映像投影を行い, その見た目を变化させる動的プロジェクションマッピングが注目を集めている. この動的プロジェクションマッピングを実現するためには, 対象物体の位置姿勢をリアルタイムかつ正確に推定する必要があり, 精度, 速度の面で優れたマーカを用いる手法が提案されている. しかしながら, マーカを付与した対象物体の作成は手間がかかり, マーカは投影された映像の見た目を悪くしてしまう. そのため, マーカを用いない位置姿勢推定手法が理想的である. 筆者らは対象物体の輪郭を用いた推定手法を提案しているものの, その推定精度や頑健性については十分に評価されていないのが現状である.

そこで本研究では, 高精度かつ頑健に位置姿勢を得ることのできるモーションキャプチャの値を ground truth とし, 筆者らの提案した手法の定量的な評価を行い, その有効性を検証する.

## 2. 従来研究

動的プロジェクションマッピング実現を目指したマーカレスな位置姿勢推定手法として, 3次元点群による位置合わせを用いるものが提案されている[1]. 3次元点群は情報量が豊富であることから, 高精度な推定が期待できるものの, 3次元計測や ICP アルゴリズムは計算時間がかかり, リアルタイムな処理が困難となる. これらの処理を高速ビジョンの仮定から高速に行う提案もされている[2]が, 高速に投影可能なプロジェクタなど特殊な機材が必要である.

一方, 筆者らは入力画像中の対象物体の輪郭と, 既知の

3D モデルの輪郭の対応付けにより位置姿勢推定を行う手法を提案している[3]. 一般に輪郭による位置姿勢推定は 2次元画像から 3次元位置姿勢を推定するため, 3次元点群などの情報量の多い情報を扱う手法に比べて精度が劣る. また, 輪郭情報は対象物体を持つユーザの手などの外乱による影響を受けやすい. そこでこの手法では, 輪郭の高速に取得できるという利点を生かし, 高速度カメラを導入することでフレーム間の差異を減らしている. 加えて, 輪郭の連続性を考慮することで, 計算時間を増やさずに対象物体の輪郭を探索する手法[4]を導入した. 以上の取り組みによって, 輪郭同士は正確に対応付けされ, 高精度かつ頑健な位置姿勢推定が実現されている.

## 3. モーションキャプチャを用いた定量評価

筆者らの提案した輪郭に基づく位置姿勢推定手法の精度および外乱に対しての頑健性の評価を行う. 提案手法は, 図 1 に示すように, 既知である 3D モデルの輪郭と, カメラ画像中の対象物体のエッジの対応付けに基づいて位置姿勢推定を行っている. 3D モデルの姿勢は前フレームの推定結果から運動予測を用いて予測され, このモデルから輪郭抽出を行い入力画像上の 2次元点へ射影する. この輪郭点から入力画像のエッジの探索を行い, この時探索した距離の和を最小化することで位置姿勢推定を行う. また, この対応付けにおいて, 隣接している輪郭点とその対応先のエッジの連続性から, 対応先が同一のエッジになるように対応点を選択することで対象物体のエッジを正確に探索する.

これらの処理は合計で約 4ms であり, 動的プロジェクションマッピング実現に必要なリアルタイム性を備えている. またこの高速性によりフレーム間での輪郭の変化が小

Yuki MORIKUBO and Naoki HASHIMOTO

さくなり、高精度な推定が可能となる。加えて運動予測や輪郭の連続性を用いた対応点探索により正確な対応付けが実現され、推定精度や頑健性が向上される。

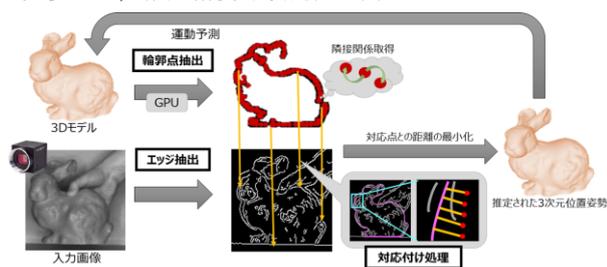


図1 提案手法のアルゴリズム

この推定手法が理想的なマーカーベースの推定精度にどの程度近づいているかを評価した。ここでは、マーカーベースの手法同様に高精度に物体の位置姿勢を推定できるモーションキャプチャを ground truth として、筆者らの提案した手法と比較することで、推定精度について定量的な評価を行った。また、対象物体を意図的に遮蔽することで、ユーザの手といった外乱に対する頑健性について定量的な評価を行った。

### 3.1 実験環境

図2に実験環境及び今回の評価で用いた対象物体を示す。対象物体にはマーカーを付与したマネキンの模型を用いる。このマーカーは対象物体の重心を正確に求めることができるように3DCGソフトウェアであるblender上で付与されている。モーションキャプチャは対象物体を動かす範囲をとり囲むように6台設置し、提案システムの推定で用いるカメラとキャリブレーションを行い同一の座標系で計測を行えるようにした。

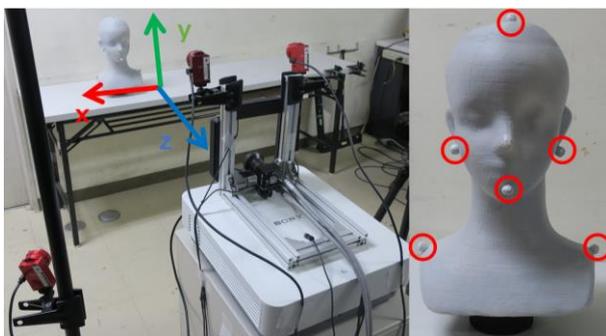


図2 実験環境と対象物体

### 3.2 推定精度の評価

提案手法の推定精度を評価するために、対象物体を手で持って動かしたときの、システムの推定値と ground truth の差を計測した。動かす方向は、奥行方向の動きの有無を考慮して、x軸およびz軸方向の並進運動とy軸およびz軸回転運動とした。各軸方向を図2に示す。並進運動では、中心から軸の方向にそれぞれ約15cm、毎秒15cmの速さで動かし、回転運動では中心からそれぞれ約45°、毎秒45°の速さで動かした。それぞれの運動について2000フレーム計測を行い、得られた推定値の ground truth との差について、平均値及び最大値を表1にまとめた。

表1 推定精度評価結果

	X位置 /mm	Y位置 /mm	Z位置 /mm	X回転 /°	Y回転 /°	Z回転 /°
X軸並進運動	4.84 (9.94)	2.28 (4.92)	4.86 (15.02)	4.32 (10.17)	1.45 (7.48)	0.51 (2.36)
Z軸並進運動	1.16 (4.21)	2.28 (4.61)	5.56 (18.75)	4.25 (10.52)	2.32 (7.00)	0.34 (1.25)
Y軸回転運動	1.70 (5.53)	2.40 (4.53)	4.06 (14.01)	2.62 (7.91)	1.64 (4.48)	0.75 (3.72)
Z軸回転運動	5.02 (11.69)	1.56 (5.21)	2.33 (7.84)	1.96 (6.65)	1.88 (6.69)	2.27 (5.75)

※ ()内は最大値

表1より提案手法は、並進成分は2~4mm程度、回転成分は1~4°程度の誤差で推定を行っていることがわかる。特に推定精度の良い推定値では1mmもしくは1°程度の誤差となっており高精度に推定できていることがわかる。しかしながら、運動方向の推定値の誤差が大きくなる傾向にあり、最大値はどの推定値も大きい。このことから、対象物体の動きが速く変異が大きい場合での推定精度に課題があることが考えられる。

誤差が大きかったx軸並進運動時のx位置およびz軸回転運動時のz回転の軌跡を図3および図4に示す。これらの結果は、運動の中心や折り返し地点で誤差が大きくなっており、このことから変異が大きい場合の精度に課題があることが考えられる。しかし、全体的に物体の移動している量に比べて誤差が小さく、精度良く推定できていることがわかる。

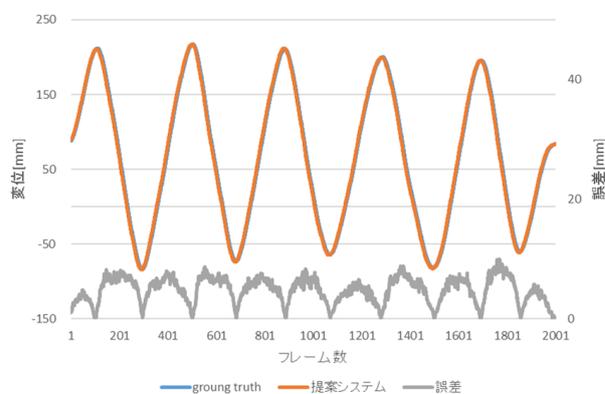


図3 x軸並進運動時のx位置

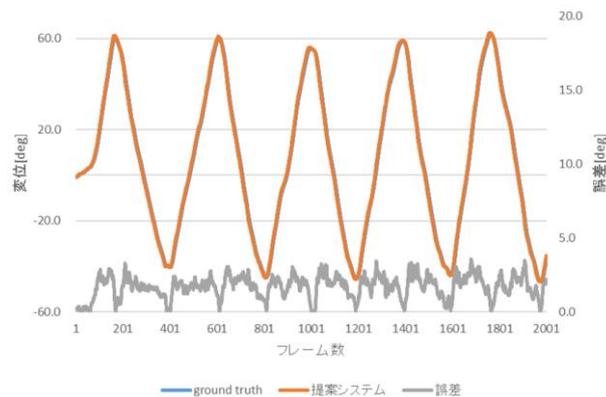


図4 z軸回転運動時のz回転

### 3.3 頑健性の評価

提案手法がどの程度の外乱まで適応可能か調べるため

に、遮蔽されている輪郭の割合と、その時の推定値と ground truth の差を測定した。また、提案の対応付け手法を導入する前の手法との比較を行った

この評価では、図5のように輪郭が隠れるようにテクスチャのない白い紙を対象物体に取り付け、y軸回転方向に3.1同様に回転運動させた。遮蔽率を変えてそれぞれ2000フレーム計測した結果を表2に示した。なお、遮蔽率は既知の3Dモデル上の輪郭点から入力画像中の対象物体の輪郭を探索した際に得られた対応点数を、輪郭点の総数で割ることで求めている。遮蔽率が15%前後の時は片手で物体を持った場合、25%前後の時は両手で物体を持った場合に対応する。



図5 遮蔽された対象物体：左から遮蔽小（約22%）、中（約28%）、大（約33%）

表2 頑健性評価結果

	遮蔽率 /%	X位置 /mm	Y位置 /mm	Z位置 /mm	X回転 /°	Y回転 /°	Z回転 /°
提案なし1 【通常状態】	13.1 (21.6)	1.46 (4.59)	2.34 (4.93)	6.66 (15.35)	2.45 (7.81)	1.66 (5.21)	0.76 (3.72)
提案なし2 【遮蔽：小】	21.8 (31.9)	3.34 (14.47)	6.57 (22.56)	4.15 (10.37)	4.86 (26.15)	5.79 (33.57)	1.19 (7.31)
提案あり1 【遮蔽：小】	23.6 (38.5)	1.57 (5.77)	2.57 (4.32)	4.04 (11.32)	1.99 (8.95)	1.75 (7.28)	0.65 (3.73)
提案あり2 【遮蔽：中】	28.5 (47.3)	1.74 (6.52)	2.62 (4.66)	4.14 (11.27)	1.67 (6.55)	1.79 (6.57)	0.53 (2.61)
提案あり3 【遮蔽：大】	33.8 (52.6)	2.06 (7.24)	2.92 (6.18)	10.90 (31.12)	2.88 (18.93)	2.59 (14.44)	1.38 (13.49)

※()内は最大値

表2の提案なし2の結果において、誤差が提案なし1より大きくなっていることから、提案した対応付け手法を用いない場合、高精度な推定を行うためには遮蔽率が約22%以下でなくてはならないと分かる。同様に提案あり3の結果から、提案手法を導入した場合に対応できる遮蔽率は平均すると34%ほどであることがわかる。しかし、提案あり2で3.1節での結果に近い推定精度が出ていることから、一時的な遮蔽であれば47%近くまでの遮蔽に耐えられることがわかる。これらの結果から、提案した対応付け手法は対象物体を持つ手といった外乱に対して頑健化が図れていることが分かる。また、提案手法はユーザが物体を両手で持って扱うには十分な頑健性であることが分かった。

図6および図7には提案あり2の運動の様子とその時の遮蔽率のグラフを示した。これらの結果から遮蔽率が40%を超えると誤差が大きくなるのがわかる。この結果からも一時的に耐えられる遮蔽率が40%ほどであることがわかる。また、平均30%ほどの遮蔽率において、外乱の影響をあまり受けず、誤差の少ない推定が行えていることが

わかる。

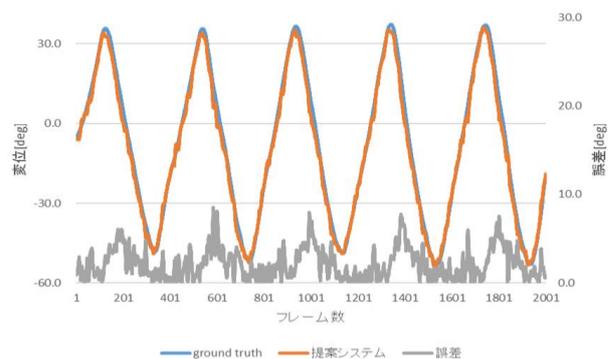


図6 提案あり2のy軸回転運動

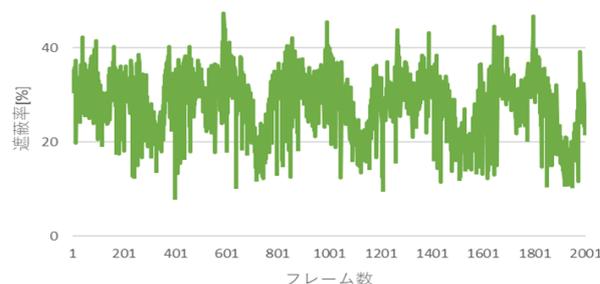


図7 提案あり2の遮蔽率

#### 4. おわりに

本研究では、一般に精度および頑健性に課題がある輪郭を用いた位置姿勢推定手法において、フレーム間の差異を小さくし対応付けを正確に行うことでこの課題を解決した筆者らの提案手法を評価した。評価結果において提案手法は、ground truth としてのモーションキャプチャの推定値に近い結果を出していた。このことから、高精度に位置姿勢推定が行えるマーカベース手法に近い推定精度が実現できていることを示すことができた。また、外乱に対して一般的な輪郭による位置姿勢推定よりも頑健であり、ユーザが手で持って扱うには十分な頑健性であることが示された。

#### 参考文献

- [1] 小林, 小泉, 橋本: “深度カメラを用いた実時間姿勢追跡に基づく動的な空間型ARの実現”, 電子情報通学会論文誌, J99-D, 3, pp. 264–272, (2015)
- [2] 田畑, 天野, 渡辺, 石川: “高速3次元トラッキングを用いたマーカレスダイナミックプロジェクションマッピング”, 信学技報, vol.117, no.391, PRMU2017-130, pp. 147–152, (2018)
- [3] N. Hashimoto, R. Koizumi and D. Kobayashi: “Dynamic Projection Mapping with a Single IR Camera”, Journal of Imaging Science and Technology, pp.020401-1-020401-24 (2016)
- [4] Y. Morikubo and N. Hashimoto: “Edge-based object tracking for dynamic projection mapping”, International Workshop on Advanced Image Technology (2018)