



# Dynamic Projection Mapping のための 輪郭に基づく反復計算によるロバストな位置姿勢推定

Robust Edge-based Iterative Posture Estimation for Dynamic Projection Mapping

森久保優輝, 橋本直己

Yuki Morikubo and Naoki HASHIMOTO

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.ucc.ac.jp)

**概要:** 本研究では, ユーザが投影対象を自由に把持して動かすことのできるプロジェクションマッピングの実現を目指す. これには, リアルタイムかつ遮蔽にロバストな 3 次元位置姿勢推定が必要であるが, 汎用的な機材により実現でき, かつ, リアルタイム性に優れた対象の輪郭を利用した研究において, 手による遮蔽の影響が大きく, 自由な把持操作が妨げられてきた. そこで本研究では, 遮蔽の影響を減らしつつ反復計算を行い, 推定精度を高めることで, 投影対象の多様な持ち方を許容する 3 次元位置姿勢推定手法を提案する.

**キーワード:** Dynamic Projection Mapping, Edge-based Object Tracking, 把持物体

## 1. はじめに

近年, プロジェクタによる映像投影で実物体の見た目を变化させる空間型拡張現実感技術が, プロジェクションマッピングと呼ばれて広く一般に普及している. プロジェクションマッピングは, ユーザに HMD 等を装着させる必要がなく, 実空間に対して CG 映像の重畳を行うことで, 現実と仮想の空間が一体となった臨場感のある演出を行うことができる. その中でも, この映像投影技術を動的環境へと拡張した Dynamic Projection Mapping (DPM) は, ユーザが手で持って動かすことのできる物体や, 自由に動き回るユーザ自身, 伸縮や歪曲などの変形をする衣服等を投影対象として想定しており, 幅広い領域での応用が期待されている.

本研究では, ユーザが把持している物体を対象とした DPM に着目する. 把持物体を対象とする DPM では, 投影映像により見た目が変化した物体を直接接触ことができ, ユーザが自由に操作することで様々な角度から眺めることができる. そのため, 製品開発におけるデザイン検討や, 医療現場や学校での教育補助等への利用が期待される. その一方で, ユーザに把持された投影対象の 3 次元位置姿勢の推定が大きな課題となっている. 多くの DPM システムでは, 投影対象や設置場所が限定されてしまうことから, 対象物体へのセンサの埋め込みや取り付けを行わず, 物体形状などのカメラによって得られる情報を基に位置姿勢推定を行っている. そのため, ユーザが投影対象を把持した場合, 手による遮蔽によって対象物体の状態を表す情報

を正確に得ることができず, 位置姿勢の推定精度が低下するため, 投影対象の持ち方や動かし方が制限されていた.

そこで本研究では, ユーザが自由に物体を把持して動かすことのできる DPM の実現を目標とし, 汎用的なカメラを用いて取得できる物体輪郭に着目した, 手による遮蔽にロバストな 3 次元位置姿勢推定手法を提案する.

## 2. 関連研究

対象物体を把持する手による遮蔽に対してロバストな 3 次元位置姿勢推定手法として, Watanabe ら[1]は, 対象物体に付与した特殊なマーカを基にした位置姿勢推定手法を提案している. この手法では, マーカが 3 点以上検出できれば推定を行うことができ, リアルタイム性にも優れているため, ユーザによる自由な操作に追従した映像投影を行うことができる. しかし, マーカを付与した対象物体の作成には手間がかかり, 付与したマーカが投影映像の見た目を損ねてしまうという問題があることから, DPM への応用には適していない.

マーカを用いずに, 対象物体が持つ情報に基づいて 3 次元位置姿勢推定を行う場合, 遮蔽領域を検出し, その領域を取り除いて姿勢推定を行う手法が多く用いられている. Held ら[2]は色情報を基にヒストグラムを作成し, 人の肌をモデル化することで手による遮蔽領域の検出を行い, 遮蔽領域を取り除いて姿勢推定を行っている. しかし, DPM では投影映像の影響を受けてしまうため, 色情報を用いたこの手法を適応することはできない. DPM に適応可能な

手法として、Zhou ら[3]は深度情報を用いて遮蔽領域を検出している。この手法では、物体形状が既知であることから、推定された物体の深度と、深度カメラによって取得できる深度を比較することで遮蔽領域を検出し、把持物体へのDPMを実現している。しかし、3次元形状復元の処理には時間がかかり、デバイスの遅延も大きいので、リアルタイム性に課題がある。

三河ら[4]は、対象物体の輪郭情報を基に、高速ビジョンにおける仮定からマスク領域を生成し、RANSACによる選択的な円検出を行うことで、DPMに向けた遮蔽にロバストな球体追跡手法を提案している。この手法は2次元画像を基に対象物体の追跡を行うため、リアルタイム性に優れているが、特殊な専用機材を必要とし、対象物体は球体や円に限定される。

### 3. 提案手法

本研究では、一般的な形状の物体を、ユーザが自由に把持して動かすことのできるDPMを実現するための、リアルタイムかつ手による遮蔽に対してロバストな3次元位置姿勢推定手法を提案する。本手法では、汎用的なカメラのみを用いて実現でき、リアルタイム性に優れた物体輪郭に基づいた位置姿勢推定手法を採用する[5]。この手法において、遮蔽による輪郭情報の欠損や、手等の対象物体以外の輪郭の影響を考慮して、対象物体の輪郭として信頼できるものを選別しながら、同一入力画像に対して反復的に姿勢推定を行うことで推定精度を高め、遮蔽に対してロバストな3次元位置姿勢推定を実現する。図1には提案手法の概要を示す。

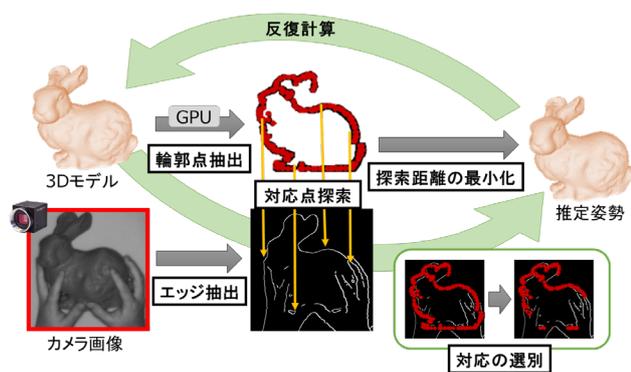


図1：輪郭に基づいた位置姿勢推

本研究における輪郭とは、物体外郭のエッジを示す。ほとんどの物体において、ある姿勢に対応した輪郭が存在しているため、その輪郭と対応する対象物体と同一形状の3Dモデルの姿勢を計算することで、2次元情報である輪郭情報から対象物体の3次元的位置姿勢を推定することができる。そこで、3Dモデルから抽出した輪郭を、カメラによって撮影された画像へ射影し、対応する対象物体の輪郭を探索する。そして、この探索距離を誤差として最小化することで位置姿勢を推定する。

本手法で用いる3Dモデルの輪郭は、前フレームに推定された位置姿勢の3Dモデルから輝度勾配を基に抽出する。また、カメラ画像の輪郭情報としては、現在のフレームにおいて撮影された画像に、Canny edge detector及びSobel filterを適用した画像を用いる。

輪郭に基づく位置姿勢推定のリアルタイム性を最大限に発揮するために、高速な撮影が行えるカメラを用いる。この撮影は、投影映像の干渉により推定精度の低下が起きないように、可視光領域を遮断するフィルタをカメラに取り付け、近赤外領域で行う。また、高速化のために、3Dモデルからの輪郭抽出はGPU上で計算を行い、3Dモデル及びカメラ画像中の輪郭抽出は並列に処理を行う。カメラ画像のエッジ抽出処理は前フレームの物体位置よりROIを設定することで、限定した領域にのみ行う。

以降の節では、輪郭に基づいた位置姿勢推定を遮蔽に対してロバストにするための手法について説明する。説明のために、3Dモデルの輪郭を構成する3次元点を、カメラ画像上の2次元点へと射影したものを輪郭点と呼ぶ。また、輪郭点と対応関係にある、カメラ画像中の対象物体輪郭上の2次元点を対応点と呼ぶ。

#### 3.1 対応点探索

基本戦略として本手法では、対応点は対象物体の同一輪郭上に存在することを利用した対応点探索手法[6]を用いる。この手法では、処理速度が非常に速い場合、フレーム間での物体の移動量が小さいことから、探索範囲は輪郭点付近のみに限定することで、対応点探索の複雑化を防いでいる。また、図2中にて青丸で示される既知の対応点から緑丸で示される対応点の候補へ方向と輪郭の方向が最も類似する対応点候補を選択し高精度な対応付けを実現している。

しかし、この手法は、対象物体の輪郭が明確に検出できている必要があり、遮蔽により正しい対応点が候補に含まれない場合には、誤った対応点を選択してしまっていた。そこで、方向の類似度が閾値以上のものだけを対応点として採用し、十分な量の対応点を得るのではなく、より正しい対応点だけを得るように拡張して、提案手法に適用した。

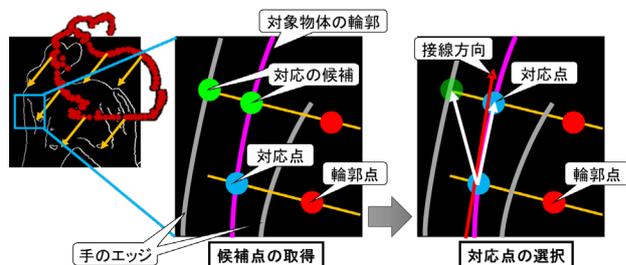


図2：対応点探索手法

#### 3.2 対応点密度による対応点の選別

3.1節の対応点探索において、対応点の候補が得られない輪郭点が存在する。これは、対象物体の輪郭が遮蔽され

てしまい、限定した探索範囲内にエッジが存在しなくなっているからである。同じく対象物体の輪郭が遮蔽されている場合、1つしか得られなかった対応点の候補は対象物体の輪郭とは限らず、遮蔽している物体のエッジであった場合には誤った対応点を選んでしまう。

そこで、対応点候補が得られなかった輪郭点付近は遮蔽されている可能性が高いという仮定から、一定領域の輪郭点における対応点尾取得の密度に基づいて、遮蔽されていない領域の信頼できる対応点を選別する。本手法では、輪郭点を10個ごとの領域に分け、60%以上の点輪郭点に対応点を取得できていない場合、その領域すべての輪郭点を後の位置姿勢推定処理に用いないようにする。この処理により、ある程度信頼できる輪郭点を取得し、次節に示す反復計算により、さらに信頼できる輪郭点を選びながら姿勢推定の精度を高めていく。

### 3.3 反復計算による対応点の選別

輪郭に基づく位置姿勢推定では、フレーム間での姿勢変化が微小であることを仮定している。しかし実際には、カメラ画像中の対象物体と同じ3Dモデルの位置姿勢を得ることができないため、輪郭同士を正確に対応付けることは難しい。また、対象を遮蔽している物体との誤対応も起きてしまい、1度の推定だけでは十分な精度が得られない。そこで、輪郭同士の対応関係を改善させながら、反復推定を行うことでより精度の高い位置姿勢推定を行い、手による遮蔽がある状況下でもロバストな推定を実現する。

本手法では、対応点の探索距離を最小化することで位置姿勢を推定しており、また、反復計算の過程において3Dモデルの姿勢は更新されカメラ画像中の物体姿勢により近いものとなる。このことから、反復計算によって輪郭点と対応点の距離は徐々に近づいていくと考えられる。そのため、より信頼できる対応点を得るために、図3に示すように、前フレームと現フレームにおける同一の輪郭点において、探索距離の増加したものは対応関係が信頼できないと判断し、推定処理から取り除く。この処理及び3.2節で信頼できないと判断された対応点は、後の反復推定でも計算に用いない。

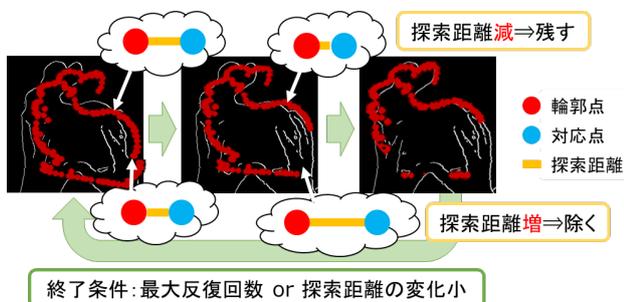


図3: 探索距離による対応点選別

以上の処理により信頼できる対応点のみを残していくことで推定精度が向上し、その結果より信頼のできる対応点を選択することができる。この反復により、対象が遮蔽

されていてもロバストに3次元位置姿勢を推定することができる。反復推定は、前フレームと現フレームの探索距離があまり変化しなくなるか、最大反復回数を越えた時に終了しカメラ画像を更新した後次の推定を行う。

## 4. 評価実験

提案手法を実装し、処理速度及び反復推定の試行回数の計測を行った。また、投影対象を自由に把持して動かした際の対応点の選別結果を描画し、確認した。実験にはCPU: Intel Core i9-9900K 3.6GHz 及び GPU: NVIDIA GeForce GTX 1080Ti, メモリ: 32GB を搭載したPCを使用した。また、高速度カメラはXIMEA MQ003MG-CM (解像度 648×488, 更新レート 500fps) を、プロジェクタはDepthQ HD (解像度 1280×720, 更新レート 120fps) を使用した。投影対象としては、把持して動かすことのできるような物体として、3Dプリンタを用いて作成したウサギ (Stanford Bunny) とサル (Suzanne) のキャラクターのモデルを使用した。

### 4.1 処理速度の測定

対象物体として Stanford Bunny を用いて、処理速度及び反復推定の試行回数の測定を行った。対象物体を両手で持ち、様々な持ち方で自由に動かしたときにおける9,000フレームの反復推定平均試行回数は3.14回であった。この時最大反復回数はプロジェクタの更新レートを下回らないように5回とし、探索距離の平均値が1.2ピクセルを下回ったとき反復推定が終了するように閾値を決定した。また、この時に行った位置姿勢推定全28,237回の平均処理速度は、表1に示す結果となった。これより、実質的な平均処理速度は4.58ms (約220fps) となり、高速な処理が実現されていることが確認できた。これはプロジェクタ更新レートである120fpsと比較するとほぼ倍の速度であり、デバイスによる遅延を予測によって補償することを考慮すると、DPMに有効な処理速度であると考えられる。

表1: 処理速度

処理内容	平均値 [ms]
輪郭点抽出・エッジ抽出	1.03
対応点探索	0.20
対応点の選別	0.16
位置姿勢推定	0.07
一回の位置姿勢推定処理合計	1.46
実質的な処理速度 (反復推定平均試行回数での処理速度)	4.58

### 4.2 輪郭点選別結果の確認

2つの対象物体について、多様な持ち方をし、自由に動かしたときの対応点選別結果を描画した結果を主観的に評価した。反復推定により最終的に選別された対応点をエッジ画像上に青色の点で描画した結果を図4(c)に示す。描

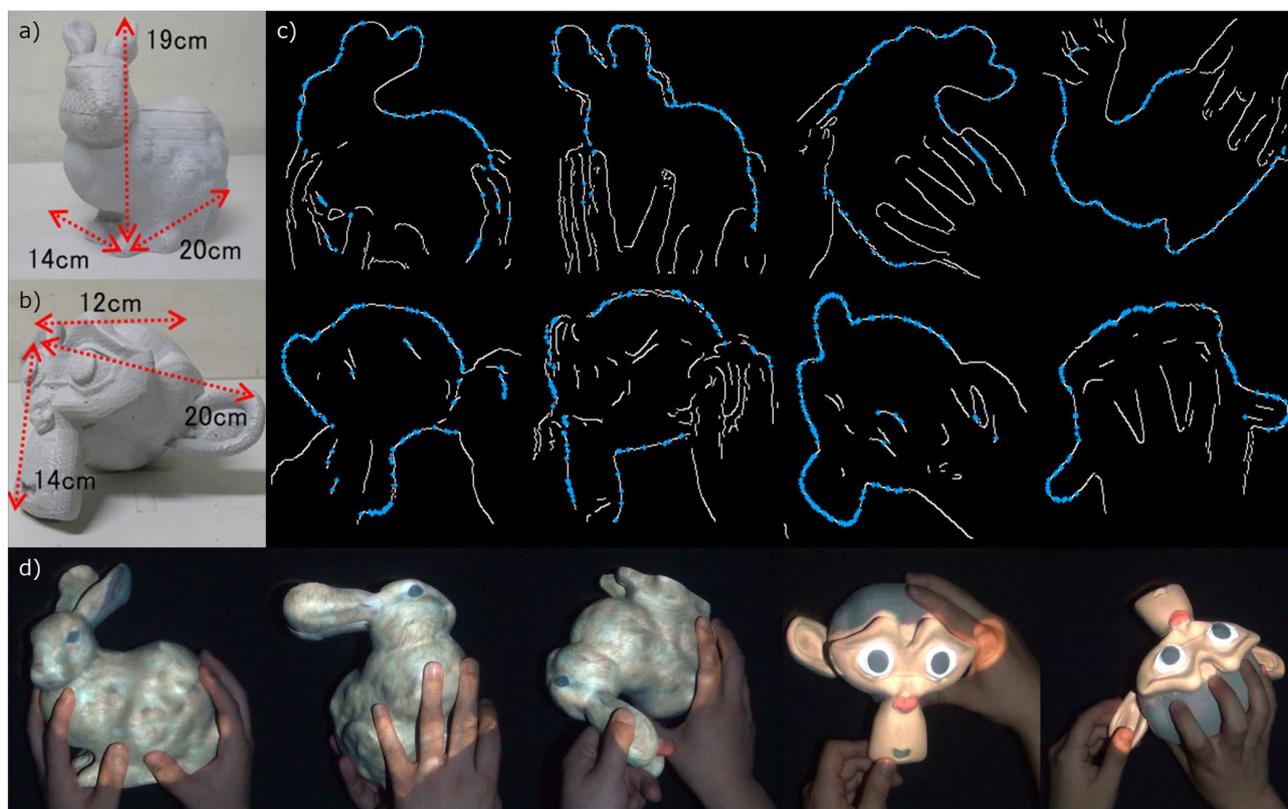


図 4 : 対応点描画結果と投影結果, (a)Stanford Bunny, (b)Suzanne, (c)選別された対応点, (d)投影結果

画結果から、手による遮蔽が起こっている部分では対応点が取得されておらず、対象物体の輪郭が明確に出ている部分では、適切に対象物体上に対応点が取得されていることが確認できた。プロジェクションマッピングという性質上、投影対象を覆い隠すような遮蔽が起きるケースはまれであり、遮蔽に対して十分頑健な位置姿勢推定が実現できていると考えられる。

この時選別された対応点の個数は、輪郭点の総数の 60～70%程度であった。このことから、輪郭点のうち少なくとも 60%程度が対応点を取得できていれば位置姿勢推定が行えると考えられる。

#### 4.3 投影実験

提案手法によって推定された位置姿勢を基に 2 つの対象物体にテクスチャを投影した。この時の結果を図 4(d)に示した。手による遮蔽がある状況下でもテクスチャが設定された位置に正しく投影されており、把持物体を対象とした DPM が実現できていることが確認できた。

### 5. おわりに

本研究では、自由に把持して動かすことのできる物体を対象とした DPM を実現するためのリアルタイムかつ遮蔽にロバストな位置姿勢推定手法を行った。今後は、遮蔽の度合いに対する推定精度を定量的に評価する予定である。また、背景や他の体操物体などによりエッジ画像が乱雑であってもロバストに姿勢推定が行えるような拡張を検討する。

謝辞 本研究は JSP 科研費 JP19H04152 の助成を受けた。

#### 参考文献

- [1] Yoshihiro Watanabe, Toshiyuki Kato and Masatoshi Ishikawa : “Extended Dot Cluster Marker for High-speed 3D Tracking in Dynamic Projection Mapping”, In Proceeding of IEEE ISMAR, pp.52-61, 2017.
- [2] Robert Held, Ankit Gupta, Brian Curless, and Maneesh Agrawala : “3D puppetry: a kinect-based interface for 3D animation”, In Proceedings of UIST, pp.423-434, 2012.
- [3] Yi Zhou, Shuangjiu Xiao, Ning Tang, Zhiyong Wei, Xu hen : “Pmomo: Projection Mapping on Movable 3D Object”, In Proceedings of the CHI Conference on Human Factors in Computing Systems, pp.781-790, 2016.
- [4] 三河祐梨, 末石智大, 石川正俊 : 動的プロジェクションマッピングに向けた輪郭情報に基づく高速球体トラッキング, 第 22 回日本バーチャルリアリティ学会大会論文集, 1E4-03, 2017.
- [5] Yuki Morikubo and Naoki Hashimoto : “Marker-less Real-Time Tracking of Texture-less 3D objects from a Monocular Image”, In proceeding of SIGGRAPH ASIA Posters No.50, 2017.
- [6] Yuki Morikubo and Naoki Hashimoto : “Edge-based object tracking for dynamic projection mapping”, International Workshop on Advanced Image Technology, 2018.