Vol.2021-CG-184 No.17
Vol.2021-DCC-29 No.17
Vol.2021-CVIM-227 No.17
2021/11/5

視覚による把持位置推定と力覚による多指ハンドの頑健な把持方策獲得

齊藤 大智[1,2,a)]　笹渕 一宏[2]　和家 尚希[2]　高松 淳[2]　小池 英樹[1]　池内 克史[2]

概要：ロボットを人間の日常生活に導入するには，人間の手を模倣した多指ハンドによるタスクを考慮した把持を可能とすることが重要である．また，位置姿勢や大きさ，形状といった物体情報が未知の環境での把持が必要である．本研究では視覚情報と力覚情報を組み合わせた新しい把持の枠組みを提案する．視覚情報である物体の写った深度画像から把持位置姿勢を推定する．把持位置姿勢の真値は人間のデモンストレーションから計算される．そして，力覚情報である指に作用する反力から多指ハンドの行動を決定することで，位置姿勢の推定誤差に適応した把持方策を実現する．把持方策学習時に用いた物体と大きさと形状が異なる物体を対象に実験を行った．その結果，把持位置姿勢の推定誤差が生じた状況であっても頑健な把持を行えることを示した．

Grasp Detection Using Vision and Robust Grasp Policy Using Haptics for Multi-fingered Hand

Daichi SAITO[1,2,a)]　Kazuhiro SASABUCHI[2]　Naoki WAKE[2]　Jun TAKAMATSU[2]
Hideki KOIKE[1]　Katsushi IKEUCHI[2]

Abstract: In order to introduce a robot into daily life, it is important to grasp with a multi-fingered hand that mimics the human hand. In addition, it is necessary to grasp in an environment where object information such as position, orientation, scale and shape is unknown. In this study, we propose a new grasp framework that combines vision and haptics. The grasp position and orientation are estimated from the depth image of the object, which is visual information. The ground truth of the grasp position and orientation is calculated from a human demonstration. Then, the grasping policy is adapted to the estimation error by determining the action of the multi-fingered hand from the haptics. We conducted the experiment using objects with different shapes from the one used for learning the grasping policy. As a result,it was shown that the system can perform robust grasping even in the situation with the estimation error.

Keywords: Grasp detection, reinforcement learning, multi-fingered hand

1. はじめに

現在，世界中で高齢化が進み家庭内での介助が望まれる．しかし，共働き世帯の増加や核家族化によって，必ずしも家庭内での介助が可能であるとは限らない．その問題を解決するために，掃除や料理のような一部の家事動作 [26] を代行する家庭用ロボットの導入が望まれる．

把持は複雑な操作を行うために必須な要素である．ロボットを人間の日常生活に導入するためには，ロボットが人間の把持を前提に作られた物体を把持できることが望ましい．そのため，人間の手を模倣した多指ハンドによる

[1] 東京工業大学情報理工学院
[2] Applied Robotics Research, Microsoft, Redmond, WA, 98052, USA
[a)] saito.d.ah@m.titech.ac.jp

Vol.2021-CG-184 No.17
Vol.2021-DCC-29 No.17
Vol.2021-CVIM-227 No.17
2021/11/5

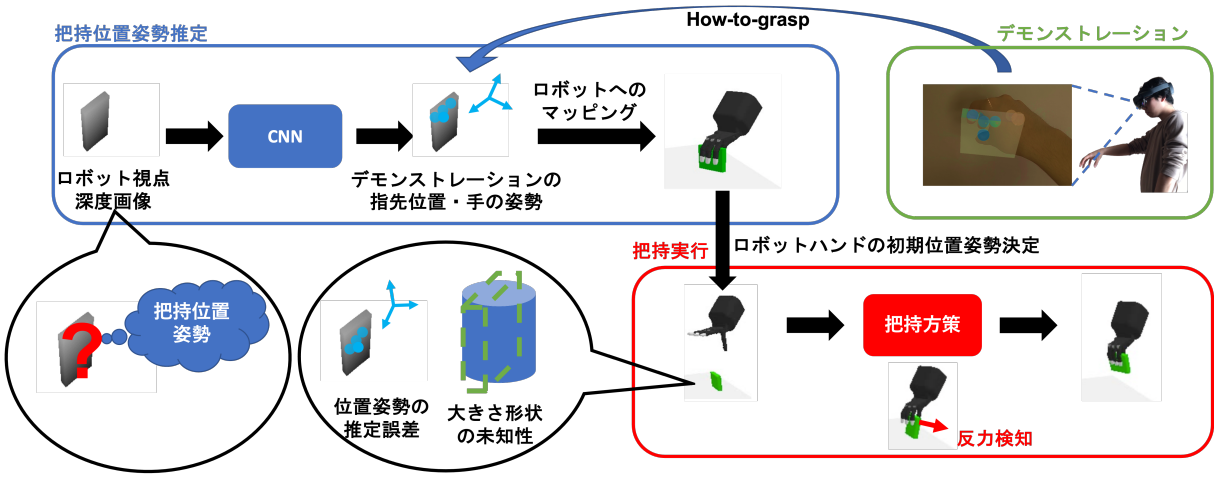

図 1 本システムの概要図

把持が期待されている．また，器用な操作を行うためには把持後のタスクを考慮した把持形態 [6] で安定把持を行う必要がある．したがって，多指ハンドによるタスクを考慮した安定把持が必要である．さらに，一般に家庭環境は多様であるため，把持する物体の位置姿勢や大きさ，形状といった物体情報が未知の環境下で把持を行えることが必要である．

既存手法による把持計画 [8][10] では，指先位置と力方向から計算される ϵ-metric [7] を用いて把持の安定性を評価し，安定把持位置を計算している．しかしながら，これらの手法ではタスクを考慮しないで把持計画をしているという問題点や，実行時には把持対象物体の三次元モデルが必要であるという問題点がある．

物体情報が未知の環境下での把持を行う手法 [14], [22] では，視覚情報である画像から把持位置姿勢を推定することで把持を行なっている．しかしながら，これらの手法は開ループ制御であり位置姿勢推定誤差や形状誤差による問題が発生する．誤差に対処するために触覚センサを用いることで把持を行う手法が存在する [1]．この手法では現時点のセンサ出力に応じて手作業で定めたルールに基づいて次時点の行動を決定することで誤差に対処する．しかしながら，これらは簡単な構造のロボットハンドであれば適用可能であるが，多指ハンドのような高次元の状態行動空間を持つ場合にはルールの設定に手間がかかるという問題点がある．

深層強化学習技術の進歩により，いくつかの研究 [2], [21] では高次元の状態行動空間を持つ多指ハンドの把持方策の学習に成功している．これらの研究では，人間のデモンストレーション [21] や把持時の人間の手と物体の接触分布 [2] といった人間の知識を用いることで，探索空間を適切に限定している．しかしながら，これらの適用範囲は物体情報が既知の環境下での把持に限られている．

そこで本研究では，タスクを考慮するために人間のデモンストレーションから計算された把持動作を初期値として与えて，その動作を元に物体との様々な接触を学習することで物体の位置姿勢や大きさ，形状が未知の環境に対処することを提案する．把持実行時には視覚情報を用いて把持位置姿勢の推定を行った後に力覚情報を用いて学習を行なった方策により把持を行うことで物体情報が未知の環境下での適応的な把持を行う (**図 1**)．力覚情報を用いた把持方策はルールベースの方策ではなく，深層強化学習によって学習される学習ベースの方策である．把持実行システムは視覚システムと力覚システムから構成されている．視覚システムでは物体の写った画像から畳み込みニューラルネットワーク (CNN) を用いることで把持時の指先位置と手の姿勢を推定する．力覚システムは人間のデモンストレーションから計算された把持時の指先位置や力方向を参考にして把持を行う．力覚情報によって物体との衝突検知や物体の姿勢情報を推定し物体の転倒を防ぐように行動することで位置姿勢誤差に対処できると考えられる．また，物体の姿勢情報を推定して行動を変化させるという方策は把持における汎用的な方策であるため，学習時の物体に対して大きさや形状がある程度変化しても適用できると考えられる．

本研究の貢献は以下の三つである．

- 人間のデモンストレーションからタスクを考慮した安定把持位置を計算する．
- 人間のデモンストレーションから計算された指先位置と力方向を参考にした報酬設計をすることで，深層強化学習による多指ハンドの頑健な把持方策獲得を実現する．
- 視覚による大まかな把持位置姿勢推定と力覚による推定誤差を許容した把持方策という二つの問題に分割することで，困難であった未知の環境での把持方策の獲得をすることを提案する．

本研究では，指先に力覚センサのついた四本指のロボッ

Vol.2021-CG-184 No.17
Vol.2021-DCC-29 No.17
Vol.2021-CVIM-227 No.17
2021/11/5

トハンド Shadow Dexterous Hand Lite [25] で把持を行う状況を想定している．また，指先が物体の上端付近にぶつかることによる転倒では物体の姿勢が著しく変化するために適応的な把持方策を行うことが強く要求されると考え，上から精密把持を行う状況を想定する．また，対象物体のみが机上に置いてあり，その物体の位置姿勢や大きさ，形状が未知である状況を想定している．

以降，2 章ではロボット把持に関する関連研究を紹介し，3 章で本システムを構成する触覚システムと視覚システムについて述べる．4 章では本システムの評価を行い，5 章で考察を述べる．最後に 6 章で結論と今後の展望を述べる．

2. 関連研究

2.1 安定把持計画

これまでに安定把持位置を計算するための様々な手法が提案されてきた [8][10]．これらは安定性の指標である ϵ-metric [7] を用いている．この指標を用いて把持候補をランク付けすることで把持位置を計算している．しかしながら，これらの手法ではタスクが考慮されない問題点や把持対象物体の三次元モデルを必要とするという問題点がある．本研究では，人間のデモンストレーションから計算された把持位置を真値として学習された CNN を用いて深度画像から把持位置を推定する．そのため，推定された把持位置はタスクを考慮した把持位置になっており，実行時には三次元モデルは必要ない．

2.2 視覚情報のみを用いた把持

深層学習技術の進歩により，画像を用いた学習ベースの把持手法が提案されている．いくつかの研究 [14], [22] では，ロボット視点の画像からロボットハンドの把持位置姿勢を推定する手法を提案している．これらは開ループ制御であるため推定誤差が把持の成功に影響を与える可能性がある．現時刻のロボット視点の画像から常にロボットハンドの把持位置姿勢を推定する手法 [28] や強化学習を用いて出力する手法 [11] といった閉ループ把持の手法も提案されている．これらの研究では視覚情報のみしか用いられていないが，本研究では力覚情報を用いた多指ハンドの把持を目標としている．

2.3 触覚情報を用いた把持

物体の把持のような環境との相互作用があるタスクでは，カメラからでは隠れて見えない部分が生じてしまい，視覚のみでは十分な情報を得ることができない．したがって，触角情報を用いることは安定把持の実現において重要である．把持の枠組みに触覚センサを統合する研究は多く行われている．例えば，手作業で定めたルールに基づき，現時点でのセンサ出力に応じて次時点の行動を決定する手法 [1] がある．これらは，簡単な構造のロボットハンドであれば適用可能であるが，多指ハンドのように高次元の状態行動空間を持つ場合にはルールの設定に手間がかかる．他には，視覚情報と触覚情報を組み合わせて把持の成功率を予測することで最適な把持を実現する研究 [4] がある．これは何度も物体に接触することで把持を行う手法で，一度の接触で把持が可能な手法と比較すると非効率である．触覚情報を用いて一度の接触で把持を行う方策を学習する手法 [16] も存在するが，これは物体情報が既知の手法である．本手法は，物体情報が未知の環境での把持を想定している．

2.4 人間の知識を組み込んだ深層強化学習による多指ハンドの把持方策獲得

多指ハンドは自由度が高く高次元の状態行動空間を持つため，探索空間が広くなってしまう．そのため，深層強化学習の適用範囲は単純な構造のロボットハンドに限られていた．そこで，人間の知識を用いることで探索空間を適切に限定する手法が提案された．例えば，人間のデモンストレーションを用いることにより方策を初期化し，タスク固有の報酬によって学習を進める模倣学習 [21] がある．また，物体と人間の手の接触分布の事前知識を用いた報酬設計をすることで深層強化学習を行う手法 [2] が存在する．これらは，物体の正確な位置姿勢が常に観測でき，物体の大きさや形状は学習時と同一であるという仮定が置かれている．これらの手法では物体の位置姿勢が観測できない場合には物体の状態を観測できる手段がないため，物体の状態に応じて行動を変えることが困難である．一方で，本手法では力覚情報により物体の状態を観測できるため物体情報が未知の環境での把持が可能である．また，物体情報が未知の環境を想定した手法として [15] が挙げられる．ロボット視点の画像から手と物体の接触点群を推定し，把持実行中に視覚を用いてそれらを常に追跡できるという仮定の元で把持を行う視覚情報のみを用いた手法である．しかしながら，把持のような環境との相互作用があるタスクでは，カメラからでは隠れて見えない部分が生じてしまうため，実環境でこの手法を適用することは困難であると考えられる．本手法では隠れてしまう問題に対して視覚情報ではなく力覚情報で対処している．

3. 手法

本研究の目標は，物体の位置姿勢や大きさ，形状が未知の環境下で把持を行うことである．本研究で提案するシステムは視覚システムと力覚システムの二つのシステムから構成されている．視覚システムでは，視覚情報である物体の写った深度画像から把持した際の指先位置と手の姿勢を推定する．力覚システムでは，視覚システムで推定された指先位置と手の姿勢からロボットハンドの初期位置姿勢

Vol.2021-CG-184 No.17
Vol.2021-DCC-29 No.17
Vol.2021-CVIM-227 No.17
2021/11/5

を決定した後に把持方策を用いて把持を実行する．その際に，力覚情報である指先への反力を用いた把持を行う．視覚システムや力覚システムは人間のデモンストレーションから計算された把持位置姿勢を用いて学習を行う．力覚情報によって物体との衝突検知や物体の姿勢情報を推定し物体の転倒を防ぐように行動することで位置姿勢誤差や大きさ，形状誤差に対処する．つまり，視覚システムでは粗い認識を行い，力覚システムでは粗い認識を許容した方策を実行する．

タスクを考慮するため，把持形態によって方策が異なる．本研究では上から精密把持をする把持形態用の方策を作成する．方策が学習時の物体と大きさと形状がある程度異なる物体を把持できることを目指す．本章では把持方策の学習手法に関して説明する．

3.1 人間のデモンストレーションの改善

人間のデモンストレーションは，我々が以前開発したシステム [23] を用いて取得する．このシステムでは，Mixed Reality Head Mounted Display である HoloLens2 を被った人間が仮想空間で物体の把持を行うことができる．人間のデモンストレーションからは把持した際の物体座標系における指先位置，手の位置姿勢が抽出された．

仮想空間では物体の重さを感じないため安定把持を達成するよりも指先を物体に接触させることを優先してしまう可能性がある．そのため，必ずしも指先位置が安定な把持位置になっているとは限らない．そこで force closure[19] に基づき指先位置の改善を行う．関節可動域内で意図した指先位置を実現するために手の姿勢が調整されていると考え，手の位置姿勢を固定した上で指先位置のみを改善した (**図 2**(a))．入力が指先位置で出力が ϵ-metric [7] である関数を遺伝的アルゴリズムを用いて最大化することで指先位置を改善した．その際に入力の指先位置と元の指先位置の距離が指定した範囲外である場合には大きな負の値を返すことで制約を加えた．

3.2 人間のデモンストレーションからロボットへのマッピング

仮想空間での把持は物体からの反力を感じないため，把持時の関節角度が現実とは異なって指の腹ではなく指の先が物体面に接触してしまうような形を取る．このような場合には点接触であるため，デモンストレーションの関節角度をそのまま用いても把持が安定しないと考えられる．そこで本手法ではデモンストレーションの指先位置と物体形状より得られる力方向から把持した際のロボットの関節角度へのマッピングを行う (図 2(b))．指先位置が指定した値に，指先の向く方向が指定した力方向と直交するように逆運動学を解くことでマッピングを行う [27]．

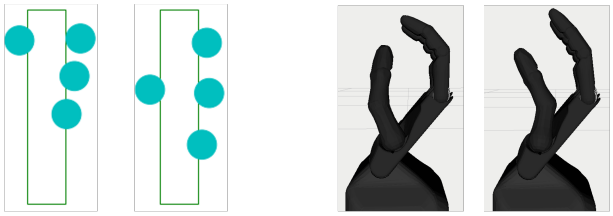

図 2 (a) 指先位置の改善 (緑線が箱の側面，青色の円が指先)．(b) 関節角度のマッピング (左が人間のデモンストレーションの関節角度，右がマッピング後の関節角度)．

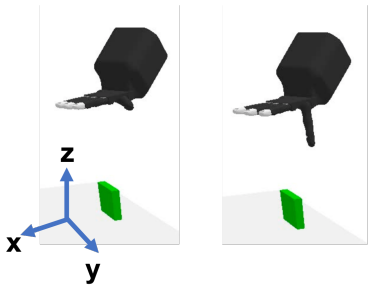

図 3 初期の手形状の図．左が全ての関節角度を 0° に設定した場合，右が対向指が存在しない方向への物体への衝突を防ぐために，親指の対立の角度と対向指の外転の角度を人間のデモンストレーションから得られた値に，その他の関節角度は 0° に設定した場合．また，本実験での座標系を示した．

3.3 参照動作生成

本手法では人間のデモンストレーションを用いて生成された参照動作を深層強化学習によって改善することで把持動作を学習する．参照動作は時刻 $t(0 \leq t \leq N)$ の手の位置 $p_t^{hand} \in \mathbb{R}^3$，手の姿勢 $q_t^{hand} \in \mathbb{R}^4$，関節角度 $j_t^{finger} \in \mathbb{R}^{16}$ から構成される．

把持した時刻を $t = N_a$ とすると，$0 \leq t \leq N_a$ の参照動作では手が把持物体に近づくようなアプローチ方向に進み，$N_a + 1 \leq t \leq N$ では把持物体を持ち上げるために手はアプローチ方向と逆方向に進む．すなわち，$p_{N_a}^{hand}$ を人間の把持した際の把持位置（ただし，前述の逆運動学の収束具合に応じて微修正を加える），p_0^{hand} を $p_{N_a}^{hand}$ から事前に定義されたアプローチ方向と逆方向に $n$cm 進んだ位置として，手は $0 \leq t \leq N_a$ で $p_t^{hand} = (1 - \frac{t}{N_a})p_0^{hand} + \frac{t}{N_a}p_{N_a}^{hand}$，$N_a + 1 \leq t \leq N$ で $p_t^{hand} = (\frac{N-t}{N-N_a-1})^2 p_{N_a}^{hand} + (\frac{t-N_a-1}{N-N_a-1})^2 p_0^{hand}$，と表せる軌道を辿る．また，$j_{N_a}^{finger}$ を人間のデモンストレーションからマッピングされた関節角度，q_0^{hand} を人間が把持した際の把持姿勢とすると，関節角度は $0 \leq t \leq N_a$ で $j_t^{finger} = (1 - \frac{t}{N_a})j_0^{finger} + \frac{t}{N_a}j_{N_a}^{finger}$，$N_a + 1 \leq t \leq N$ で $j_t^{finger} = j_{N_a}^{finger}$，手の姿勢は $0 \leq t \leq N$ で $q_t^{hand} = q_0^{hand}$ で表せる軌道を辿る．

なお，j_0^{finger} の設計は使用するロボットハンドの特性

Vol.2021-CG-184 No.17
Vol.2021-DCC-29 No.17
Vol.2021-CVIM-227 No.17
2021/11/5

によるが，本研究で使用するロボットハンドの場合には，対向指が存在しない方向への物体への衝突を防ぐために，アプローチ開始位置では親指の対立の角度と対向指の外転の角度を人間のデモンストレーションから得られた値として，その他の関節角度は 0° に設定した (**図 3**).

3.4 視覚情報による把持位置姿勢推定

物体の写った画像を I，把持位置姿勢を G とする．本章の目的は，物体の写った画像から把持位置姿勢を推定する関数 $F : I \to G$ を学習することである．

物体表面のテクスチャの違いの影響を受けないように，入力画像 $I \in \mathbb{R}^{224\times224}$ には深度画像を用いた．出力 $G \in \mathbb{R}^{19}$ は人間のデモンストレーションから得られた把持時の指先位置 $p \in \mathbb{R}^{15}$ と手の姿勢 $q \in \mathbb{R}^{4}$ である．手の姿勢はクォータニオンで表現されている．また，出力はカメラ座標系に変換されており，指先位置は前処理後の値になっている．深度画像は物理シミュレータである Bullet Physics SDK [5] を用いて収集した．カメラから物体までの距離やカメラの姿勢はランダム化されている．

本研究では畳み込みニューラルネットワークを用いて把持位置姿勢推定を行なった．ネットワーク構造は四層の畳み込み層と三層の全結合層である．カーネルサイズが $16\times3\times3$，$32\times3\times3$，$64\times3\times3$，$128\times3\times3$ の畳み込み層の後に，出力が 1024，256，19 の全結合層に入力された．畳み込み層の後には 2×2 の MaxPooling が適用される．活性化関数は最終層以外は ReLU [18]，最終層には線形関数を用いた．ネットワークの重みは He の初期化 [9] により初期化された．(1) 式で表される損失関数 L は指先位置 p の誤差に関する L2 ノルムと手の姿勢 q の誤差に関する L2 ノルムの和である [12]．ある姿勢を表現するクォータニオンは二つ存在するため（q と $-q$ は同じ姿勢を表現している），真値と推定値の和と差の小さい方を損失関数に加えた．学習アルゴリズムには Adam [13] を用いた．また，過学習を防ぐために validation の精度が一定期間上がらなくなったら学習を打ち切る Early Stopping を行なった．

$$L = \|\hat{p} - p\|_2^2 + min(\|\hat{q} - \frac{q}{\|q\|}\|_2^2, \|\hat{q} + \frac{q}{\|q\|}\|_2^2) \quad (1)$$

3.5 力覚情報を組み込んだ把持方策

多指ハンドの把持方策学習には深層強化学習を用いる．エージェントは指先位置と力覚情報は取得できるが，物体の位置姿勢や大きさ，形状は取得できないという環境を想定している．

定式化　把持方策学習をマルコフ決定過程として定式化する．マルコフ決定過程は状態空間 $\mathcal{S}$，行動空間 $\mathcal{A}$，状態遷移 $\mathcal{T} : \mathcal{S}\times\mathcal{A} \to \mathcal{S}$，初期状態分布 ρ_0，価値関数 $r : \mathcal{S}\times\mathcal{A} \to \mathbb{R}$ を持つ．我々の目標は，割引率を $\gamma \in [0, 1)$ として累積報酬和 $J(\pi) = \mathbb{E}_\pi\left[\sum_{t=0}^{T-1}\gamma^t r\left(s_t, a_t\right)\right]$ を最大化する方策 $\pi(a|s)$ を求めることである．ここで，T はエピソードの長さ，$s_0 \sim p_0$，$a_t \sim \pi(s_t)$，$s_{t+1} = \mathcal{T}\left(s_t, a_t\right)$ である．

状態　状態 s_t は $s_t = \{p_t^{finger}, f_t^{finger}, t\}$ で定義される．ここで，$p_t^{finger} \in \mathbb{R}^{12}$ は各指先位置，$f_t^{finger} \in \mathbb{R}^{12}$ は各指の反力の方向である．

行動　行動 a_t は $a_t = \{jd_t^{finger}, d_t, g_t, w_t\}$ で定義される．本手法では人間のデモンストレーションから生成された参照動作との差分を行動としている．jd_t^{finger} は各指の関節角度の差分，$d_t \in \mathbb{R}$ は手の位置のアプローチ方向成分の差分，$g_t \in \mathbb{R}$ は手の位置のアプローチ方向と手首軸との外積方向成分の差分，$w_t \in \mathbb{R}$ は手首軸周りの回転の差分である (**図 4**).

また，各指の探索空間を狭めるため，本研究で使用するロボットハンドにおいて，人差し指・中指・薬指の三本の指を連動させる．これは，人間が人差し指・中指・薬指を用いた精密把持をする場合には，これらを連動させて動かしていると考えられるからである．つまり，人差し指，中指と薬指の外転以外の関節角度の差分を $jd_t^{index} \in \mathbb{R}^3$，$jd_t^{middle} \in \mathbb{R}^3, jd_t^{ring} \in \mathbb{R}^3$ とすると，$jd_t^{middle} = jd_t^{ring} = jd_t^{index}$ となる．また，人差し指と薬指は中指に対してほぼ対称の位置に存在すると考え，人差し指の外転の角度が d である時に，中指の関節角度は 0，薬指の関節角度は $-d$ とした．この関節角度は安定把持の形成に影響しないため学習は行わずに常に固定値とした．

報酬　報酬は (2) 式で表されるアプローチ時の報酬 $r_t^{approach}$ と，(3) 式で表される持ち上げ時の報酬 r_t^{lift} の二つある．多様な行動を促すことで様々な状態に対処できるようにするために指定した指のみに関する報酬設計をした．アプローチ時には指定した指の指先位置 p_t^{finger} と参照動作の指先位置 p_t^{ref} とのユークリッド距離 $dist_t^{approach} = \|p_t^{ref} - p_t^{finger}\|_2$ が設定した閾値 α_t より小さければ正の報酬を，大きければ負の報酬を与えた．持ち上げ時には指定した指にかかる力の方向 f_t^{finger} と指定した力の方向とのコサイン距離 $dist_t^{lift} = f_t^{ref} \cdot f_t^{finger}$ が設定した閾値 α^{lift} より小さければ正の報酬を，大きければ負の報酬を与える．多様な行動を学習させるために把持動作開始時には広い範囲を探索し，徐々に探索範囲を狭めて把持時には参照動作の指先位置付近を探索する (ファネル戦略) ように動的な閾値を設計した．また，不自然な動きの学習を避けるために Early termination を行う [20]．この閾値の範囲にもファネル戦略を適用した．アプローチ時には $dist_t^{approach} \leq \beta_t$，持ち上げ時には $dist_t^{lift} \leq \beta_t$ の時に大きな負の報酬 r_{et} を与えてエピソードを終了させる．また，物体が倒れた時にも同様の処理を行なう．

学習はシミュレータ内で行われた．強化学習フレームワークには Microsoft 社が提供する Project Bonsai [17] を使用した．学習アルゴリズムには Actor-Critic 法で方策オ

Vol.2021-CG-184 No.17
Vol.2021-DCC-29 No.17
Vol.2021-CVIM-227 No.17
2021/11/5

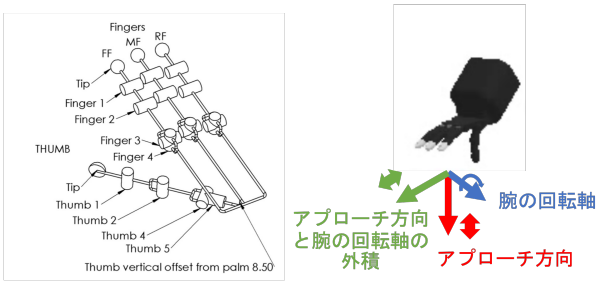

図 4 ロボットハンドの構造 [25] とアプローチ方向を重力方向とした時の手の並進と回転の方向を表した図．関節の行動は Thumb1,2,3,4 と FF の Finger1,2,3 の行動が学習される．

ン手法である Proximal Policy Optimization [24] を用いた．

$$r_t^{approach} = log(1 - (dist_t^{approach} - \alpha_t)) \quad (2)$$

$$r_t^{lift} = log(1 - (dist_t^{lift} - \alpha_t)) \quad (3)$$

4. 実験

実験では 2 つのことを調査する．視覚システムの学習時の物体と学習時とは大きさと形状が異なる物体での把持位置姿勢推定誤差を求めることで視覚システムの汎用性を調査する．次に，視覚システムによる推定誤差が発生する環境下で力覚システムが学習時の物体や大きさと形状が学習時とは異なる物体を把持できるかどうかを調査する．実験はシミュレータ内で行われた．本実験の座標系は図 3 に示した通りである．

物体は (4) 式で表される superquadrics [3] を用いて表現した．a_1，a_2，a_3 は物体の x 軸方向の大きさ，y 軸方向の大きさ，z 軸方向の大きさを表している．ϵ_1，ϵ_2 は xz 平面における形状，xy 平面における形状を表している．

$$\left(\left(\frac{x}{a_1}\right)^{\frac{2}{\epsilon_2}} + \left(\frac{y}{a_2}\right)^{\frac{2}{\epsilon_2}}\right)^{\frac{\epsilon_2}{\epsilon_1}} + \left(\frac{z}{a_3}\right)^{\frac{2}{\epsilon_1}} = 1 \quad (4)$$

本実験では，比較的倒れやすい x 軸方向の長さ 2cm，y 軸方向の長さ 10cm，z 軸方向の長さ 10cm の直方体 ($a_1 = 0.01, a_2 = 0.05, a_3 = 0.05, \epsilon_1 \simeq 0, \epsilon_2 \simeq 0$) を上から精密把持する人間のデモンストレーションが取得された．

実行時には，アプローチ方向は重力方向に設定された．把持開始位置は把持位置からアプローチ方向と逆方向に $n = 20$cm 進んだ位置とした．また，エピソードは時刻 $N = 13$ までとしアプローチは時刻 $N_a = 6$ までとした．親指の位置や反力方向に関しての報酬が得られ，$f_t^{ref} = (-1, 0, 0)$ とした．報酬の動的閾値はアプローチ時は $\alpha_t^{approach} = 0.08 - 0.01 \times t$，持ち上げ時は常に同じ値とし $\alpha^{lift} = 0.03$ とした．Episode Termination の閾値はアプローチ時の $\beta_t = 0.1 - 0.01 \times t$，持ち上げ時の $\beta_t = 0.4 - 0.01 \times (t - (N - N_a))$，$r_{et} = -100$ とした．物体の重心が開始地点から 1cm 下がったら物体が倒れたとした．

(a) 指先位置の平均絶対誤差 (cm)

	学習時に用いた物体群	学習時に用いていない物体群
親指	0.270 ± 0.155	0.284 ± 0.167
人差し指	0.273 ± 0.165	0.293 ± 0.183
中指	0.271 ± 0.163	0.273 ± 0.169
薬指	0.268 ± 0.163	0.284 ± 0.172
小指	0.286 ± 0.172	0.310 ± 0.185

(b) 手の姿勢の平均絶対誤差 (°)

	学習時に用いた物体群	学習時に用いていない物体群
yaw	0.938 ± 1.048	1.065 ± 1.313
pitch	1.450 ± 1.102	1.473 ± 1.175
roll	0.480 ± 0.466	0.492 ± 0.503

表 1 学習時に用いた物体と用いていない物体での推定誤差の比較

4.1 視覚システム

$a_1 = 0.01, a_2 = 0.05, a_3 = 0.05, \epsilon_1 \simeq 0, \epsilon_2 \simeq 0$ 付近の $0.01 \leq a_1 \leq 0.05, 0.06 \leq a_2 \leq 0.20, 0.06 \leq a_3 \leq 0.16, \epsilon_1 \simeq 0, 0 < \epsilon_2 \leq 2$ の範囲の物体で視覚システムの学習・評価を行なった．ϵ_1 によって形状の変化した物体は上からの精密把持の対象外として値は固定されている．a_1 の範囲の端点，a_2, a_3, ϵ_2 の範囲の端点と中点の組み合わせの物体 54 個とデモンストレーションで用いた物体と同じ大きさと形状，もしくは異なる形状の $a_1 = 0.01, a_2 = 0.05, a_3 = 0.05, \epsilon_2 \simeq 0, \epsilon_2 = 1, 2$ の 3 個の物体を学習に用いた．物体からカメラまでの距離やカメラの姿勢がランダム化されて収集された約 25 万枚の深度画像を使って学習された．評価では a_1 を 0.01 ずつ動かした値，a_2, a_3 を 0.02 ずつ動かした値，ϵ_2 を 0.5 ずつ動かした値の組み合わせの物体と学習時に用いた物体の 1233 個の物体を用いた．各物体に対して深度画像が 63 枚取得された．

学習時に用いた物体群と学習時に用いていない物体群で評価した結果が**表 1** である．評価指標として位置誤差には平均絶対誤差 $\|\hat{p} - p\|$ を用いた．姿勢誤差はクォータニオンの差 $\hat{q}q^{-1}$ に対して yaw-pitch-roll の絶対値の平均値を求めた．どちらの群の精度も同程度の精度であり，内挿ができていることが分かった．本実験では親指位置を用いて初期位置を計算する．親指の位置誤差は学習時に用いた物体と用いていない物体を合わせたデータセットで x 軸方向に -0.058 ± 0.199cm，y 軸方向に -0.092 ± 0.117cm，y 軸方向に -0.032 ± 0.224cm であった．また，姿勢誤差は yaw 軸周りに $-0.123 \pm 1.677°$，pitch 軸周りに $-0.869 \pm 1.672°$，roll 軸周りに $-0.008 \pm 0.703°$ であった．本実験では 2σ 区間までの誤差が発生するとして，力覚システムの学習ではこの範囲内での誤差を発生させた．

Vol.2021-CG-184 No.17
Vol.2021-DCC-29 No.17
Vol.2021-CVIM-227 No.17
2021/11/5

表 2 推定誤差が発生する環境での把持成功率の比較．ベースラインは参照動作である．

	ベースライン	本手法
成功率	68.8%	100%

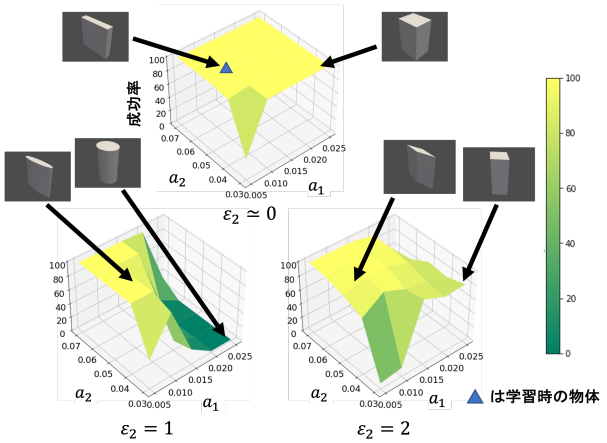

図 5 本手法を用いて上面の大きさと形状 (a_1, a_2, ϵ_2) が異なる物体を把持した場合の把持成功率．

4.2 力覚システム

視覚システムによる推定誤差が発生する環境下で力覚システムが学習時の物体や大きさや形状が学習時とは異なる物体を把持できるかどうかを検証した．エピソード終了時に物体を把持していれば成功と定義した．力覚が頑健な把持において重要であることを検証するために，ベースラインである開ループ制御方策を用いた場合との比較を行なった．ベースラインではアプローチ時は参照動作を行い，物体を把持して落とさないように持ち上げ時は MC 関節の外転と CP 関節の対立以外の関節角度を参照動作よりも 0.2rad 曲げることで把持を行う．

視覚システムによって 2σ 区間の端点の誤差が発生したとして位置誤差と姿勢誤差の組み合わせの $2^6 = 64$ 回の把持を行なった．学習時の物体で把持を行なった結果が**表 2** である．ベースラインでは成功率が 68.8% であったが，本手法では成功率が 100% であった．

前の実験と同じ位置姿勢誤差を発生させた環境下で，本手法によって学習時と大きさと形状の異なる物体の把持を行なった結果が**図 5** である．物体上面の大きさと形状が変化するように superquadrics のパラメータ a_1, a_2, ϵ_2 を変化させた．形状 ϵ_2 が学習時の物体と同じ場合では大きさが異なっていても位置姿勢誤差に頑健であることが分かった．一方で，形状が異なる場合には大きさが変化すると位置姿勢誤差への頑健性が著しく低下することが分かった．

5. 考察

学習した方策がどのような行動を取ることで把持を行なっているのかを調査するために，時刻 $t = 5$ において人差し指のみで物体に触れた場合の把持データを複数取得

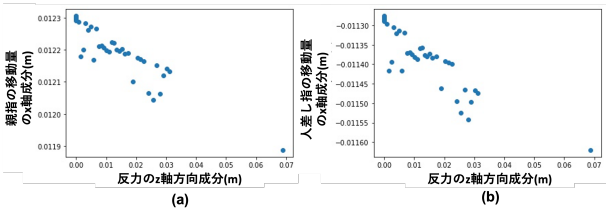

図 6 反力の z 軸成分と指先の移動量の x 軸成分の関係．(a) 親指の移動量の x 軸成分，(b) 人差し指の移動量の x 軸成分．

し，人差し指にかかる反力の方向の z 軸成分と指の移動量の x 軸成分との関係を図示した (**図 6**)．z 軸成分が大きくなるほど，物体はより倒れていることになる．指の移動量が小さくなるほど指が x 軸負方向に移動することになる．人差し指は物体が倒れるほど物体を押し，親指は物体から離れる方向へ移動する傾向があった．物体の転倒を防ぐためには人差し指を物体から離して親指を物体へ近づけるように学習するはずであるが，今回の場合には予想と反した学習をした．それでもベースラインと比較して把持成功率が高く物体のスケール形状が変化しても把持が可能であった理由は，物体と衝突して転倒しないようなアプローチを学習したからだと考えれる．力覚によって物体の状態を推定できるため頑健な把持を行うには力覚を組み込んだ把持の枠組みが重要であると考えられたが，今回の場合にはアプローチの学習の影響の方が大きいように見える．一方で，力覚を入れた恩恵として，人差し指が物体を押して親指が後ろに下がることで物体の姿勢を整えるような行動ができたことが挙げられる．これはタスクの実行のために非常に重要な行動である．

大きさと形状が同時に変化した場合に失敗してしまう例として，把持後に指が物体面に沿って滑ってしまう場合が挙げられる．物体が厚くなる場合や物体が狭くなる場合に物体面の傾きが増加するために失敗しやすくなる．これは把持方策の学習時には傾きがない物体面でしか学習していないために対処できなかったと考えられる．これは複数形状の物体を学習させることで対処可能であると考えられる．もしくは，学習する物体を傾いた面を持つ物体にすることで，物体が転倒しかけている場合への方策と傾いた面を持つ場合への方策の両方が学習できるため，より広い形状への対処が可能になる可能性もあると考えられる．

6. おわりに

本研究では視覚情報と力覚情報を用いた新しい把持の枠組みを提案した．実験の結果，本手法は把持成功率でベースラインを上回った．学習した把持方策は物体の転倒を防ぐようにアプローチをし，力覚情報を用いて物体の状態を推定することで物体の姿勢のずれを修正していることがわかった．これはタスクの実行にとって重要である．物体の大きさの変化には頑健であった一方で形状の変化には弱

かった．本研究は家庭環境のような物体情報が未知の環境下に把持操作が可能なロボットを導入するための第一歩になると考えている．今後の展望としては，力覚の効果が出る状況のさらなる調査，単一の把持方策が把持可能な物体の範囲を最大化するために学習すべき物体の大きさや形状の調査が挙げられる．

参考文献

[1] Al-Gaifi, R. M., Müller, V. and Elkmann, N.: Reactive grasping using high-resolution tactile sensors, *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*, IEEE, pp. 463–468 (2020).
[2] Añazco, E. V., Lopez, P. R., Park, N., Oh, J., Ryu, G., Al-antari, M. A. and Kim, T.-S.: Natural object manipulation using anthropomorphic robotic hand through deep reinforcement learning and deep grasping probability network, *Applied Intelligence*, Vol. 51, No. 2, pp. 1041–1055 (2021).
[3] Barr, A. H.: Superquadrics and angle-preserving transformations, *IEEE Computer graphics and Applications*, Vol. 1, No. 1, pp. 11–23 (1981).
[4] Calandra, R., Owens, A., Jayaraman, D., Lin, J., Yuan, W., Malik, J., Adelson, E. H. and Levine, S.: More than a feeling: Learning to grasp and regrasp using vision and touch, *IEEE Robotics and Automation Letters*, Vol. 3, No. 4, pp. 3300–3307 (2018).
[5] Coumans, E. and Bai, Y.: PyBullet, a Python module for physics simulation for games, robotics and machine learning, `http://pybullet.org` (2016–2021).
[6] Feix, T., Romero, J., Schmiedmayer, H.-B., Dollar, A. M. and Kragic, D.: The grasp taxonomy of human grasp types, *IEEE Transactions on human-machine systems*, Vol. 46, No. 1, pp. 66–77 (2015).
[7] Ferrari, C. and Canny, J. F.: Planning optimal grasps., *ICRA*, Vol. 3, No. 4, p. 6 (1992).
[8] Goldfeder, C., Allen, P. K., Lackner, C. and Pelossof, R.: Grasp planning via decomposition trees, *Proceedings 2007 IEEE International Conference on Robotics and Automation*, IEEE, pp. 4679–4684 (2007).
[9] He, K., Zhang, X., Ren, S. and Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034 (2015).
[10] Huebner, K., Ruthotto, S. and Kragic, D.: Minimum volume bounding box decomposition for shape approximation in robot grasping, *2008 IEEE International Conference on Robotics and Automation*, IEEE, pp. 1628–1633 (2008).
[11] Kalashnikov, D., Irpan, A., Pastor, P., Ibarz, J., Herzog, A., Jang, E., Quillen, D., Holly, E., Kalakrishnan, M., Vanhoucke, V. et al.: Qt-opt: Scalable deep reinforcement learning for vision-based robotic manipulation, *arXiv preprint arXiv:1806.10293* (2018).
[12] Kendall, A., Grimes, M. and Cipolla, R.: Posenet: A convolutional network for real-time 6-dof camera relocalization, *Proceedings of the IEEE international conference on computer vision*, pp. 2938–2946 (2015).
[13] Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980* (2014).
[14] Lenz, I., Lee, H. and Saxena, A.: Deep learning for detecting robotic grasps, *The International Journal of Robotics Research*, Vol. 34, No. 4-5, pp. 705–724 (2015).
[15] Mandikal, P. and Grauman, K.: Learning Dexterous Grasping with Object-Centric Visual Affordances, *IEEE International Conference on Robotics and Automation (ICRA)* (2021).
[16] Merzić, H., Bogdanović, M., Kappler, D., Righetti, L. and Bohg, J.: Leveraging contact forces for learning to grasp, *2019 International Conference on Robotics and Automation (ICRA)*, IEEE, pp. 3615–3621 (2019).
[17] Microsoft: Project Bonsai, `https://azure.microsoft.com/en-us/services/project-bonsai/`.
[18] Nair, V. and Hinton, G. E.: Rectified linear units improve restricted boltzmann machines, *Icml* (2010).
[19] Nguyen, V.-D.: Constructing force-closure grasps, *The International Journal of Robotics Research*, Vol. 7, No. 3, pp. 3–16 (1988).
[20] Peng, X. B., Abbeel, P., Levine, S. and van de Panne, M.: Deepmimic: Example-guided deep reinforcement learning of physics-based character skills, *ACM Transactions on Graphics (TOG)*, Vol. 37, No. 4, pp. 1–14 (2018).
[21] Rajeswaran, A., Kumar, V., Gupta, A., Vezzani, G., Schulman, J., Todorov, E. and Levine, S.: Learning Complex Dexterous Manipulation with Deep Reinforcement Learning and Demonstrations, *Proceedings of Robotics: Science and Systems (RSS)* (2018).
[22] Redmon, J. and Angelova, A.: Real-time grasp detection using convolutional neural networks, *2015 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, pp. 1316–1322 (2015).
[23] Saito, D., Wake, N., Sasabuchi, K., Koike, H. and Ikeuchi, K.: Contact Web Status Presentation for Freehand Grasping in MR-based Robot-teaching, *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 167–171 (2021).
[24] Schulman, J., Wolski, F., Dhariwal, P., Radford, A. and Klimov, O.: Proximal policy optimization algorithms, *arXiv preprint arXiv:1707.06347* (2017).
[25] Shadow Robot Company: Shadow Dexterous Hand Lite, `https://www.shadowrobot.com/products/dexterous-hand/`.
[26] Smarr, C.-A., Mitzner, T. L., Beer, J. M., Prakash, A., Chen, T. L., Kemp, C. C. and Rogers, W. A.: Domestic robots for older adults: attitudes, preferences, and potential, *International journal of social robotics*, Vol. 6, No. 2, pp. 229–247 (2014).
[27] Starke, S., Hendrich, N., Krupke, D. and Zhang, J.: Evolutionary multi-objective inverse kinematics on highly articulated and humanoid robots, *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, pp. 6959–6966 (2017).
[28] Viereck, U., Pas, A., Saenko, K. and Platt, R.: Learning a visuomotor controller for real world robotic grasping using simulated depth images, *Conference on Robot Learning*, PMLR, pp. 291–300 (2017).