



強化学習を用いたハンドリダイレクション手法 の予備的検討

加藤諒¹⁾, 松本啓吾²⁾, 鳴海拓志²⁾, 葛岡英明²⁾

1) 東京大学 大学院情報理工学系研究科 (〒 113-0033 東京都文京区本郷 7-3-1, r-kato@mi.t.u-tokyo.ac.jp)

2) 東京大学 大学院情報理工学系研究科 (〒 113-0033 東京都文京区本郷 7-3-1, matsumoto, narumi, kuzuoka@cyber.t.u-tokyo.ac.jp)

概要: 実空間でのユーザの手の動きとバーチャル空間でのバーチャルハンドの動きの対応関係を操作することで、実物体とは異なる位置や形状のバーチャルオブジェクトに触れる体験を可能にするリダイレクテッドハンドと呼ばれる手法が提案されている。本研究では、ユーザの実際の手の動きとバーチャルハンドの動きの対応関係を視触覚的破綻を負の報酬とした強化学習によって学習する手法を提案し、予備的な検証を行った。

キーワード: リダイレクテッドハンド, 強化学習, クロス・マルチモーダル, 感覚・知覚

1. はじめに

低解像度の実物体を用いて高解像度のバーチャルオブジェクトに触れる体験を可能にする技術であるパッシブハプティクスとよばれる手法が提案されている。パッシブハプティクスは、複雑な機構を必要とせずにバーチャルオブジェクトに触れる体験を可能にするが、実物体とバーチャルオブジェクトとで形状や位置・姿勢の整合性を一定程度保つ必要があり、触力覚提示可能なバーチャルオブジェクトが限定される。この問題に対して、視触覚間相互作用を用いることでパッシブハプティクスを拡張したリダイレクテッドハンド(ハンドリダイレクションとも呼ばれる)やハプティックリターゲティングが提案されている。リダイレクテッドハンドでは実物体とは形状や位置・姿勢が異なるバーチャルオブジェクトを視覚提示し、実空間におけるユーザの手の動きとバーチャル空間におけるバーチャルハンドの動きの対応関係をわずかに操作することで、実物体とは異なる形状や位置・姿勢のバーチャルオブジェクトに触れる体験を可能にしている。

このようにリダイレクテッドハンドを用いることにより触れることのできるバーチャルオブジェクトの幅は広がるが、従来のリダイレクテッドハンドでは、実空間とバーチャル空間での手の動きの対応関係であるマッピングモデルは制作者によって個別に構築されており、多様な実物体とバーチャルオブジェクトの組み合わせに対するマッピングモデルを構築することは高コストになることが予想される。

この課題を解決するため、多様な実物体とバーチャルオブジェクトとのマッピングモデルを自動生成する強化学習を用いた汎用マッピングモデルの構築を目指す。本報告ではその前段階として、実物体とバーチャルオブジェクトに1点で触れる際の2次元汎用マッピングモデルを構築し、予備的検証を行う。

2. 関連研究

2.1 リダイレクテッドハンド

リダイレクテッドハンドとは実際の手の動きと形状とはわずかに異なるバーチャルハンドの動きや形状を視覚提示することで、実物体とは異なるバーチャルオブジェクトに触れる体験を可能とする手法である [1]。伴らは、指で物体をなぞる際にディスプレイで視覚提示した物体の形状に対応するようにバーチャルハンドの位置を変化させることで、形状知覚が操作可能であることを明らかにした [3]。

2.2 ハンドハプティックリターゲティング

手を伸ばして物体を掴む際に、実空間では一つの物体に触れているにも関わらず、ユーザの手の位置を変換することで、バーチャル空間内で異なる位置に存在する複数のオブジェクトに触れていると知覚させることができるハプティックリターゲティング (Haptic Retargeting) が Azmandian らによって提案されている [2]。ハプティックリターゲティングでは、バーチャル空間上のバーチャルハンドを操作することで実空間で物体に触れている手をバーチャル空間においてもオブジェクトに触れているようにする Body Warping, バーチャル空間の座標系を操作しバーチャル空間のオブジェクトと実空間の物体の位置を揃える World Warping の2つの手法を組み合わせた Hybrid Warping という手法が用いられている。

3. 強化学習を用いた2次元汎用マッピングモデル

本研究では実空間とバーチャル空間の動きの対応関係であるマッピングモデルを図1に示すように、実空間の空間情報・バーチャル空間の空間情報・実空間での手の動きを入力、バーチャル空間での手の動きを出力とした関数と捉える。この関数の最適化を行うため視触覚的破綻と視覚体性感覚間のずれの2つの指標を負の報酬として捉えた強化学習(図1)を行う。視触覚的破綻とは実空間とバーチャル空

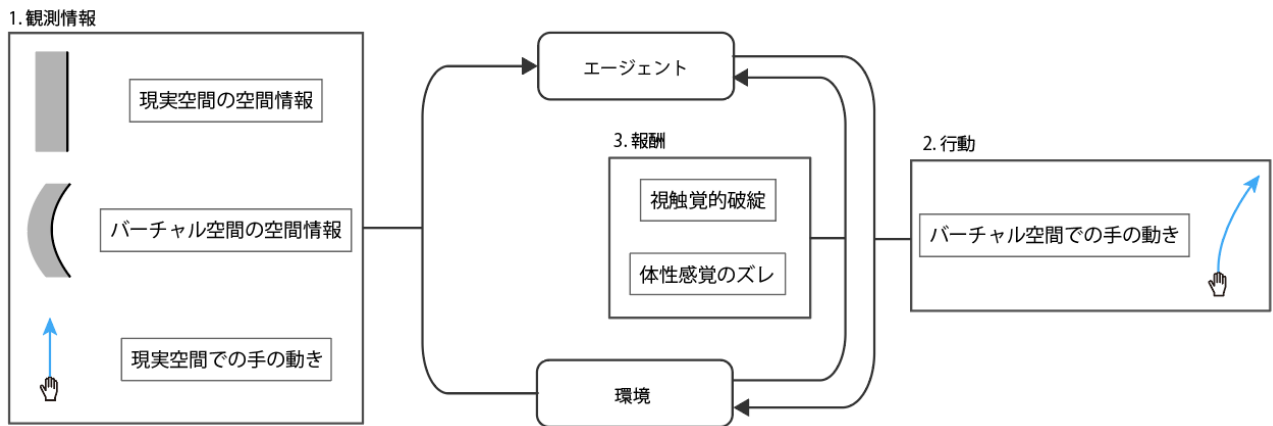


図 1: 本研究で提案する強化学習を用いた 2 次元汎用マッピングモデルの概観

間のどちらか一方のみで手が物体に接触している状況と定義する。また、視覚体性感覚間のずれとはバーチャルハンドの動きについての視覚入力と実際の手の動きにより生じる体性感覚のずれと定義する。ここで、マッピングモデルの出力（行動）は手の動き（速度）であり、その値は連続的なものであるため、強化学習のアルゴリズムとして連続的な行動空間も扱うことのできる Proximal Policy Optimization (PPO) [4] を利用する。

また、学習に用いる実空間とバーチャル空間を模したシミュレーション環境を設計した。学習の効率化と、現実空間とバーチャル空間の形状が異なる場合にも適切なマッピングができるようにするために、出来る限り多様な環境を生み出すことのできるシミュレータを設計する必要がある。視触覚的破綻に関して有意義な学習を行うことができるシミュレーションは、境界条件でのシミュレーションであるため、境界を主な想定状況として学習を行う。

4. 学習環境設計

4.1 シミュレーション環境

学習の際、マッピングモデルの入力となる実空間の空間情報、バーチャル空間の空間情報、実空間での手の動きをシミュレーションにより用意する必要がある。実空間とバーチャル空間の物体形状が異なる場合にも適切なマッピングができるよう、多様な実物体およびバーチャルオブジェクトの生成と多様な手の動きを生成するシミュレーション環境の構築を行う。

4.1.1 実物体とバーチャルオブジェクトのシミュレーション

シミュレーション環境上における実物体とバーチャルオブジェクトとして、図 2 に示すような外面と内面からなるオブジェクトを生成した。

外面は凹な形状、内面は凸な形状を表現しており凹面・凸面双方のオブジェクトに触れる体験を再現することができる。また、これらは図 2 のように大きさ・角度・曲率を操作でき、学習のエピソードごとにこれらの値がランダムに変化する。なお、オブジェクトの大きさ・角度・曲率の値が取

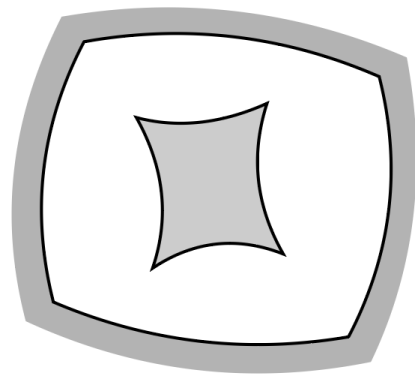


図 2: シミュレーションで生成される実物体およびバーチャルオブジェクト

表 1: 本研究におけるカリキュラムラーニング

レベル	大きさ	角度	曲率	レベルアップに必要な報酬
1	1±0.1	0°	1	-
2	1±0.2	±10°	1	5
3	1±0.4	±20°	1±0.2	8

りうる範囲は学習効率を最適化するためカリキュラムラーニングに基づき表 1 に示すように強化学習の進行度に応じて段階的に引き上げた。これらのゲインは先行研究で報告されている閾値をもとに設定した [5, 1]。

4.2 実空間での手の動きのシミュレーション

本研究では、簡素化のため実物体およびバーチャルオブジェクトに 1 点で触れる状況を想定してシミュレータを設計する。このオブジェクトに 1 点で触れるという運動をシ

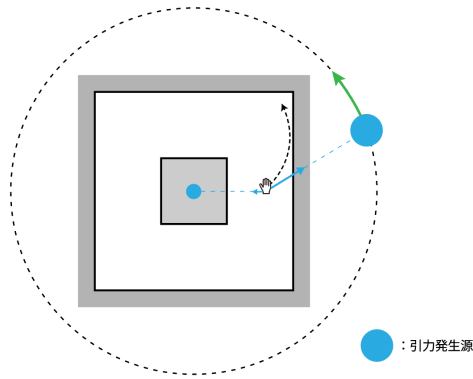


図 3: 手の動きのシミュレーション

ミュレートするために、シミュレータ上でオブジェクトの外側を公転するような引力発生源と内側の中心に置いた引力発生源を用意し、2つの引力発生源の強さと外側の引力発生源の公転速度をランダムに切り替えることで手の動きを生成する。

手は2つの引力発生源に引っ張られることで動くが、2つの引力発生源から生じる引力の強さはランダムに決定しているため、多くの場合両者は同一の値にはならない。外側からの引力の方が強い場合、手は外面に沿って移動し、内側からの引力の方が強い場合、手は内面に沿って移動する。そのため、本シミュレータはオブジェクトへのさまざまな触れ方を再現することができると考えられる。なお、このシミュレーション上の実空間では手に衝突判定を設けることで実物体を通り抜けないようにしている。

4.3 モデルの入出力

モデルで扱う入出力はすべて正規化のプロセスを経てから扱う。空間情報は $[0, 1]$ 、動きの情報は $[-1, 1]$ の範囲となるよう正規化を行う。

4.3.1 入力

本研究でのマッピングモデルの入力は実空間の空間情報、バーチャル空間の空間情報、実空間での手の動きである。実空間の空間情報は、手から周囲の物体への距離を $360/k^\circ$ ごとにまとめた長さ k のベクトル情報として扱う。本研究では空間情報の解像度とモデルのパラメータ数のバランスを考えて $k = 360$ とした。同様にバーチャル空間の空間情報についてはバーチャルハンドから周囲の物体への距離のベクトル情報として扱う。このように、周囲の物体への距離を 1° ごとにまとめたベクトル d^R 、 d^V をそれぞれ実空間、バーチャル空間の空間情報として扱った。実空間での手の動きは、手の2次元の並進速度と1次元の回転速度を合わせたスクリュウベクトル w^R として扱う。

4.3.2 出力

出力はバーチャル空間でのバーチャルハンドの動きである。これは入力の実空間の手の動きと同様に、バーチャルハンドから見た2次元の並進速度と1次元の回転速度を合わせたスクリュウベクトルとして扱う。

表 2: モデルの入出力

表現される情報	入出力	値範囲
実空間の空間情報	入力	$d^R \in [0, 1]^{360}$
バーチャル空間の空間情報	入力	$d^V \in [0, 1]^{360}$
実空間の手の動き	入力	$w^R \in [-1, 1]^3$
バーチャル空間の手の動き	出力	$w^V \in [-1, 1]^3$

4.4 報酬

視触覚的破綻のないマッピングモデルの獲得のため、視触覚的破綻がないほど多くの報酬が得られるように、破綻を数値化し負の報酬として扱う。また、実空間における手の動きとバーチャルハンドの運動のずれに起因する視覚体性感覚間の違和感を低減させるために、運動のずれに関する報酬も設計に組み込む。

4.4.1 視触覚的破綻に関する報酬

視触覚的破綻に関する報酬としては離散的な報酬と連続的な報酬を選択した。

離散的な報酬 ユーザが視触覚的破綻に気づくのは、実空間では手が物体に触れているのにもかかわらずバーチャル空間ではバーチャルオブジェクトに触れていない、もしくは、実空間では手が物体に触れていないのにもかかわらずバーチャル空間ではバーチャルオブジェクト触れている状態と想定される。離散的な報酬では、ユーザが視触覚的破綻に気づくと想定される状態のフレームごとに一定の値の負の報酬を与える。

連続的な報酬 離散的な報酬では視触覚的破綻が生じていないときは報酬が与えられないことになり十分な学習機会が得られないことが考えられる。この問題に対して、物体もしくはバーチャルオブジェクトの近傍におけるずれを報酬とした連続的な報酬を提案した。連続的な報酬は、毎フレームごとに与えられる視触覚的破綻に関する報酬 r_h をモデルの入出力で定義した実空間の空間情報 d^R 、バーチャル空間の空間情報 d^V を用いて、

$$r_h \propto \sum_{i=0}^{k-1} (1 - \max(d_i^R, d_i^V, \epsilon) / \max(\min(d_i^R, d_i^V), \epsilon)) \quad (1)$$

とした。これにより、物体への距離が近いほど両空間における距離の差が大きく負の報酬として反映される。ただし、 d_i^R 、 d_i^V がどちらも非常に小さい場合、式??だと、報酬が負の方向に非常に大きな値を取る可能性があるため、 ϵ を用いている。ただし ϵ は十分小さい数である。

4.4.2 視覚体性感覚間の運動のずれに関する報酬

実空間の動きの情報 w^R 、バーチャル空間の動きの情報 w^V を用いて、毎フレームごとに与えられる運動のずれに関

する報酬 r_c を,

$$r_c \propto -\|w^R - w^V\| \quad (2)$$

とした。

5. 2次元汎用マッピングモデルの予備的検証

本研究では、シミュレーション環境を用いて2次元汎用マッピングモデルの予備的な性能評価を行った。評価項目としてはヒューマンファクタとの関連が小さい視触覚的破綻を選択した。検証に用いるモデルとして、離散的な報酬を用いて学習したマッピングモデル、連続的な報酬を用いて学習したマッピングモデルの3種類を用いた。また、コントロール条件として恒等変換モデルを用いた。離散的な報酬を用いて学習したマッピングモデルと連続的な報酬を用いて学習したマッピングモデルは4.1節で提案したシミュレーション環境を用いて400万ステップ（フレーム）学習させた。視触覚的破綻の評価では、シミュレーション環境を用いて実物体の位置・角度・曲率と手の動きを固定し、バーチャルオブジェクトの位置・角度・曲率を変化させることで視触覚的破綻が生じる割合を測定した。具体的には、3秒かけて実物体に接近し、3秒かけて実物体に沿う手の動きを1セットとし、バーチャルオブジェクトをセットごとに変化させて計10セット行い、視触覚的破綻が生じるフレームの割合を計測した。評価の際にはバーチャルオブジェクトはモデル間で共通したセットを用いた。

表3に恒等変換モデル、離散的な報酬を用いて学習したモデル、連続的な報酬を用いて学習したモデルにおける視触覚的破綻が生じた割合を示す。

表 3: モデルの評価実験結果

モデル	視触覚的破綻が生じない割合 [%]
恒等変換	43.2
離散的な報酬	26.7
連続的な報酬	21.5

連続的な報酬を用いて学習したモデルの方が離散的な報酬を用いて学習したモデルより視触覚的破綻が生じなかった割合が小さかった。この結果は、汎用的なマッピングモデルの構築において、離散的な報酬設計と比べて情報量の大きい連続的な報酬設計で、獲得できるマッピングモデルの性能を高めることができること示唆している。他方、連続的な報酬を用いて学習したモデルであっても21.5%の時間で視触覚的破綻が生じていた。実際に、学習に用いたシミュレータでマッピングモデルを動かすと、視触覚的破綻が所々に生じていることが確認できた。特に、実空間では手が実物体に接しているにもかかわらず、バーチャル空

間ではバーチャルハンドがバーチャルオブジェクトに接していないという視触覚的破綻がみられた。

6. まとめ

本研究では強化学習を用いたリダイレクテッドハンドにおける2次元汎用マッピングモデルの構築および予備的評価をおこなった。2次元汎用マッピングモデルの構築にあたって、シミュレーション学習環境の作成、モデルの入出力の選定を行った。さらに報酬設計として実際の視触覚体験を用いた離散的な報酬とより情報量の大きい連続的な報酬の2種類を選択した。予備的評価からは連続的な報酬を用いて学習したモデルが離散的な報酬を用いて学習したモデルよりも視触覚的破綻を減少させることが示唆された。一方で、連続的な報酬を用いて学習したモデルであっても視触覚的破綻の割合は21.5%と大きかった。特に、実空間では手が実物体に接しているにもかかわらず、バーチャル空間ではバーチャルハンドがバーチャルオブジェクトに接していない視触覚的破綻がみられた。こうした、視触覚的破綻を減少させるためには、学習の手本となるデータ（エキスパートデータ）を用意し、強化学習に組み込むことが考えられる。また、提案モデルについてユーザスタディを行うことで視触覚的破綻に加えて、視覚体性感覚間のずれについても評価を行うことを検討している。

謝辞 本研究はJSPS科研費JP18J21379, JP19H04149の助成を受けたものです。

参考文献

- [1] Bergström, J., Mottelson, A., Knibbe, J. (2019, October). Resized grasping in vr: Estimating thresholds for object discrimination. In Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (pp. 1175-1183).
- [2] Azmandian, M., Hancock, M., Benko, H., Ofek, E., Wilson, A. D. (2016, May). Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences. In Proceedings of the 2016 chi conference on human factors in computing systems (pp. 1968-1979).
- [3] 伴祐樹, 鳴海拓志, 谷川智洋, 廣瀬通孝. (2012). 手の動きの空間変調による形状知覚操作 (特集) 3次元インタラクション. 日本バーチャルリアリティ学会論文誌, 17(4), 457-468.
- [4] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O. (2017). Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.
- [5] Cheng, L. P., Ofek, E., Holz, C., Benko, H., Wilson, A. D. (2017, May). Sparse haptic proxy: Touch feedback in virtual environments using a general passive prop. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 3718-3728).