

強化学習による学習型マイクロ流体システムの
開発と新規液中微粒子操作技術への応用

山梨大学大学院
医工農学総合教育部
博士課程学位論文

2023年3月

阿部 岳晃

目次

序章 緒論.....	1
第1章 マイクロバルブへの強化学習の適用と動作シーケンスの学習.....	5
1.1 緒言.....	5
1.2 理論、実験方法.....	7
1.2.1 デバイスの構造.....	7
1.2.2 強化学習の適用.....	8
1.2.3 デバイスの作製.....	11
1.2.4 学習システムの構築.....	12
1.2.5 流量のサンプリング.....	13
1.3 実験結果および考察.....	14
1.3.1 動的計画法による動作シーケンスの獲得.....	14
1.3.2 Q学習による動作シーケンスの獲得.....	19
1.3.3 Q学習による流量調整動作の獲得.....	20
1.3.4 強化学習による環境変化への適応性の実証.....	22
1.4 結言.....	25
第2章 強化学習を利用した1次元微粒子操作技術の開発.....	26
2.1 緒言.....	26
2.2 理論、実験方法.....	27
2.2.1 強化学習の適用.....	27
2.2.2 粒子操作シミュレーションおよび実験環境.....	28
2.3 実験結果および考察.....	29
2.4 結言.....	31
第3章 Vacuum-assisted Micromolding を用いたリフロープロセス.....	32
3.1 緒言.....	32
3.2 理論、実験方法.....	33
3.2.1 EVA ホットメルト接着剤のパターニングとリフロープロセス.....	33

3.2.2	モールドとデバイスの作製	34
3.2.3	流路断面とパターニング能力の評価	35
3.2.4	バルブシール性能とポンプ流量の測定	36
3.3	実験結果および考察	37
3.3.1	形状再現性評価	37
3.3.2	パターニング能力評価	41
3.3.3	バルブおよびポンプの性能評価	42
3.4	結言	44
第4章	2次元微粒子操作技術の開発	45
4.1	緒言	45
4.2	理論、実験方法	46
4.2.1	デバイスの構造と強化学習の適用	46
4.2.2	粒子操作シミュレーションおよび実験環境	47
4.2.3	CFD と 2D 物理エンジンを利用した液中粒子挙動の予測	47
4.2.3	デバイスの製造	48
4.3	実験結果および考察	50
4.3.1	予測能力の評価	50
4.3.2	粒子操作の実証	52
4.4	結言	55
第5章	深層学習を利用した粒子操作シミュレータと粒子の並列操作	56
5.1	緒言	56
5.2	理論、実験方法	57
5.2.1	デバイスの構造と強化学習の適用	57
5.2.2	強化学習アルゴリズムの適用	58
5.2.3	デバイスの作製	58
5.2.4	学習システム	59
5.2.5	シミュレータの構築	60
5.3	実験結果および考察	62
5.3.1	シミュレータの評価	62
5.3.2	単一粒子操作	63
5.3.3	粒子 2 個並列操作	67

5.4 結言.....	70
第6章 総括.....	71
参考文献.....	74
謝辞.....	83

序章 緒論

物体の精密な操作技術は、マイクロ・ナノメートルスケールにおける現象の解明に不可欠である[1]。微小スケールの現象の研究は、細胞生物学[2-5]、分子生物学[6-8]、医療診断[9]、創薬[10, 11]などの発展に貢献してきた。これまでに様々な micromanipulation system が開発されており、その操作原理は、光、磁気、電場、音響波、流体力学など、多岐に渡る。

微粒子操作技術の中で最も利用される手法は、光ピンセットとして知られる光学ベースの手法である[12, 13]。これは、集光されたレーザーの force field を利用して粒子を操作する技術で、細胞や微生物[14, 15]などの生体組織の操作にも用いられる。この手法は、高い空間的精度と被操作物質の多様性をもつが、高価な光学装置が要求される。また、レーザー光により、生体分子および組織へダメージを与える可能性がある[15]。磁気ベースの手法は、磁性を持つ粒子を、外部からの磁場を用いて制御する[6, 16]。操作対象が磁性を保つ必要があるため、磁性を持たない粒子に対しては磁気ラベリングが必要となる。電気的手法は、操作対象の粒子や周辺の流体への電圧の印加による静電的な力で粒子を操作するものである[17]。高精度な操作と並列化が可能であるが、操作は対象の分極率に依存し、低伝導率の媒体が必要で、細胞関連技術への応用が制限される[1, 18]。音響的手法は操作対象の存在する領域の外側から振動を与え、溶媒の圧力分布を制御することで粒子を操作する[19-21]。音波の発生には、基本的に interdigitated transducers (IDT) が用いられるが、この製造は多段階にわたり、デバイス製造の高コスト化が問題となる[22]。一方、流路内の流体の流れのみで粒子を操作する手法も提案されている。流路に配置されたピラー周辺に発生する流れ場を制御する手法[23]や、等角度で交差する流路の流入出を制御することで粒子を操作する手法[24, 25]などがある。このシステムは流体力学の理論をベースとして実現されるため、システムの設計と制御には流路構造に応じた数理モデルが必要となり、所望のシステムの作製には専門的知識が必要となる。よって、低コストで並列化可能なシステム、生体物質への適合性、容易な設計と制御を実現できるプラットフォームは実現されていない。

マイクロ流体システムは、微小流路を用いることで、サンプルの削減、処理の自動化、反応の高速化などの利点がある[26-28]。マイクロ流路内の流体の駆動および流れの制御は、流路構造の他に、外部ポンプ[29]や重力[30, 31]、マイクロバルブ[32, 33]を利用して実現されることが一般的である。また、micro electro mechanical systems (MEMS) の技術を応用したデバイスの作製手法により、大規模集積化によるハイスループットシステムが実現できる[34-36]。

これらの技術は、粒子操作技術にも応用される[1, 37]。しかし、複雑な機能を実現するためには、ポンプなどを多数使用し、正確に制御する必要があり、システム全体の大型化や高コスト化が課題となる。また、基本的に、システムを構成する多数の流体制御用の機能要素は、単位操作に一对一で対応する[34, 38]。このような設計手法では、様々なタスクに対応するシステムを開発することは困難である[39, 40]。さらに、システムを安定化・最適化するた

めには、設計者やシステム運用者がタスクに応じて無数のバルブの挙動をカスタマイズすることが求められる。これは、システムを熟知していないユーザーにとっては現実的でない場合がある。そのため、粒子操作のような複雑な機能を実現するためには、新たなアプローチが求められる。

近年、マイクロ流体システムと機械学習の統合が進んでおり、特にディープラーニングの応用が進んでいる[41]。Zhen らは、2017 年以降の知的マイクロ流体に関する研究を 3 つの分野に分類し、約半数がバイオテクノロジー関連、3 分の 1 がチップ設計・最適化、残りが化学分野であると報告した[42]。例えば、バイオテクノロジーにおける細胞の分類・解析技術[43-47]、液滴の形状予測[48, 49]、チャンネル構造設計[50]などが報告されている。さらに、化学の分野では、マイクロリアクタにおける化学反応の最適化[51, 52]や物質検出[53]が主な研究テーマとなっている。ここで、マイクロ流体システムへの機械学習の応用の方向性として、強化学習による制御が提案されている。強化学習は、連続的な意思決定問題に対する最適な方針を試行錯誤によって学習する機械学習手法の一種である[54]。近年、深層学習との統合により実現される、深層強化学習のパフォーマンスの高さが実証され、研究事例が増加している[55-62]。代表的な例として、ビデオゲーム[63-66]や囲碁[67, 68]で人間を超えるスキルを学習した AI がある。また、ロボットの制御[69-76]などにも用いられる。

マイクロ流体工学の分野では、デジタルマイクロ流体デバイスの微小液滴制御[77]やマイクロチャンネル構造の最適化[78]に強化学習が適用されている。Dressler らは、強化学習をシリンジポンプの制御に適用することで、超安定な液滴生成を実現した[79]。この研究は、強化学習により、環境変化への適応性を実現できることを証明した。このような研究は、強化学習の適用が連続的なタスクにおける性能向上につながることを証明した。また、強化学習は、設計者や使用者が細かく設定を行うことなくシステムの挙動を柔軟に切り替えられるような、より汎用的なマイクロ流体システムの実現に貢献できる可能性がある。

上記のコンセプトを実現するために有望な機能要素として、マイクロ流体工学で最も一般的に用いられるアクチュエータである、マイクロバルブが挙げられる[80, 81]。これは、流れのオン・オフのスイッチとして機能するだけでなく、集積された複数のマイクロバルブを制御することで、単純な流路の切り替え[82]、タスクモードの動的切り替え[83]、マイクロミキシング[84]、マイクロポンプ[85]など、様々な流体機能を実現できる。この機構に対してより高度な制御技術を適用することにより、これまで複雑なハードウェアを必要としていた、粒子制御技術などの高度な動作を、オンチップアクチュエータのみで実現できる可能性がある。加えて、強化学習の学習能力により、より柔軟で汎用的なマイクロ流体システムを実現する可能性を持つ重要な試みであると考えられる。

本論文では、オンチップマイクロバルブと強化学習の特性に着目し、これらを利用した制御技術の確立と、本システムを応用した新しい微粒子操作技術の構築を目標とする。まず、マイクロバルブの強化学習による制御を実証したあと、粒子操作システムの構築と、シミュレータの開発による学習の効率化、実環境における粒子操作の実証を行う。

第1章では、マイクロ流体システムへの強化学習の統合を実証する第一歩として、マイクロバルブを複数集積した機能モジュールであるマイクロペリスタルティックポンプを対象に、強化学習の適用性を検討した。具体的には、最大流量を達成するための最適な動作シーケンスを決定した。マイクロペリスタルティックポンプの動作はマルコフ性を満たすと仮定し、一連の動作シーケンスにおいて、ある状態から最適な次の状態を決定する意思決定問題に対して、強化学習を適用した。まず、実験システムの環境変化が軽微であると仮定し、エージェントによる探索を必要としない方法を採用した。ポンプの動作を Markov Decision Process (MDP)に基づいてモデル化・定式化した。その後、ポンプ、顕微鏡、強化学習の計算処理用コンピュータからなる学習システムを構築した。このシステムを用いて、2つの状態の間を遷移する際にダイアフラムが発生する流量をサンプリングした。流量が最大となるような動作シーケンスを得るために、サンプリングした流量を報酬とした。また、得られた動作シーケンスと従来手法の流量を比較し、本手法を評価した。その後、環境を探索することで学習を行う、Q学習の適用が行われた。Q学習では、報酬関数を未知として、環境内を試行錯誤で探索し、行動決定のための評価関数を学習する。さらに、より高度なタスクとして、ポンプの流量調整の学習を行った。これらのタスクに加え、強化学習を適用することによってシステムの環境変化への対応能力を示した。

第2章では、本システムの粒子操作技術への応用を行った。強化学習は、特徴量を設定することで、システムの目的を設定する。これを利用し、前章と同じハードウェアを用いて、システムの目的を単純な1次元粒子操作とすることで、マイクロバルブを用いた新しい粒子操作システムが実現できることを示した。実環境とシミュレータでの学習を行い、大規模化を視野に入れた学習環境の評価が行われた。

第3章では、マイクロバルブ構造の作製技術として、基板上に熱可塑性樹脂をパターンニングする技術を提案した。これは、マイクロ流体デバイスの構造材料として一般的に使用されている Polydimethylsiloxane (PDMS)製の流路に、溶融した樹脂を減圧下で注入し、流路構造をパターンニングする技術である。まず、加熱時間を変えて加熱を行い、流路の断面形状の変化と形状の再現性を調べた。次に、直角流路を用いて、マイクロ流体システムで必要とされる複雑な曲線構造の作製能力を評価した。さらに、サイズの異なるバルブを作製し、シール性などの特性を評価した。

実用的な観点から、粒子操作システムでは2次元平面上での操作が必要である。よって、第4章では、マイクロバルブと強化学習を用いた、2次元の粒子操作技術を実証した。強化学習は試行錯誤によって最適なポリシーを獲得するため、実環境のみでの学習は膨大なコストを要する。そのため、学習用のシミュレータを用いた。流体中の粒子操作タスクをシミュレートするため、バルブ動作に対する粒子の挙動を予測する手法を開発し、シミュレータに組み込んだ。学習前に Computational Fluid Dynamics (CFD)ソフトウェアを用いてデバイス内部の速度データを予測し、このデータを粒子の速度と仮定して2次元物理エンジンに入力することにより、マイクロ流路内の粒子の動きを予測する方法を提案した。本手法による

予測結果を、実測した粒子運動と比較し、シミュレータの精度を確認した。その後、シミュレータを用いた2次元粒子操作の学習が行われた。さらに、学習済みモデルを用いて実環境における実験を行うことで、提案手法を実証した。

シミュレータは基本的に、物理の理論ベースで構築されるが、デバイスや操作対象の形状に合わせて一つ一つ数理モデルを作成することは、システムの応用範囲を制限する。この煩雑さを克服するため、第5章では、バルブ動作に対する粒子の運動の関係をニューラルネットワークで予測する。このネットワークは実際のデバイスからランダムサンプリングされた学習用データを用いた教師あり学習で実現される。シミュレータの予測精度が確認された後、強化学習による粒子の操作の学習が行われた。学習したモデルは実環境にデプロイされ、粒子の2次元平面上での粒子操作の実証が行われた。

第6章では、提案した手法についての総括と、今後の展望について述べる。

第1章 マイクロバルブへの強化学習の適用と動作シーケンスの学習

1.1 緒言

本章では、マイクロ流体システムと強化学習を統合した粒子操作技術の開発に先立ち、マイクロ流体システムへの強化学習の適用を試みた結果を報告する。

マイクロ流体システムの流体制御には、マイクロバルブと呼ばれる微小アクチュエータがよく用いられる。これは、空気圧や圧電素子の変形を駆動力として、ポリマー薄膜を変形し流路を閉塞する機構である図 1.1(a)。これと同時に、薄膜の変形により薄膜上部の流体が周辺へと移動する。本機構は構造が単純であるため、マイクロ流体デバイスの製造で一般的に用いられるソフトリソグラフィの技術により、数 cm のデバイス上に数百のバルブを実装することも可能である。この特性により、バルブを複数個集積し、連動させることで、図 1.1(b)のように流体を一定方向に輸送するポンプとして利用されることがある。このバルブの動作制御は基本的にシステムの設計者や使用者によって経験的に決定されるか、これまでに提案された動作シーケンスを用いることが多い(図 1.2)。本章では、デバイス上に実装されたマイクロバルブの動作制御に対する強化学習の適用手法を確立する。また、強化学習が、これまで経験的に決定されてきたマイクロペリスタルティックポンプの効果的な制御手法となることを実証する。

まず、3つのバルブから構成されるマイクロペリスタルティックポンプをデバイス上に構成し、顕微鏡、カメラ、学習処理用のコンピュータ、空気圧制御系から成る学習プラットフォームを構築した。次に、マイクロペリスタルティックポンプの動作に対して動的計画法による動作獲得と Q 学習による学習手法の2つの手法を適用した。強化学習適用のため、マイクロポンプの動作に対してマルコフ性を仮定し、マルコフ決定過程として定式化を行い、動的計画法を用いて最適な動作の学習を試みた。

ここで、ポンプで発生する流量を強化学習における報酬と定義することでポンピング動作の獲得を図った。また、ポンプの動作で発生する流量をすべての動作パターンでサンプリングすることにより、報酬関数がモデリングされ、コンピュータ上で学習が行われた。その後、モデルフリーの学習手法である Q 学習の適用が行われた。ここでは、報酬関数は未知として、システムがマイクロポンプの状態行動空間を試行錯誤で探索することで行動決定に用いる評価関数の最適化が行われる。サンプリングされた流量をもとに作成した報酬関数を用いてコンピュータ上での学習と、実環境での学習を行い、結果を比較した。

より高度な制御として、ポンプの流量を所望の値に調整する動作の獲得を試みた。ここまでの学習では流量が最大となるような行動の学習を行ったが、これを所望の流量と現在の平均流量の偏差を報酬とすることで流量の調整動作を獲得できることを示した。

さらに、Q 学習は環境内を探索し、現在の評価関数における行動価値が最大の行動を選択する。これにより、探索中の環境の変化に対してパフォーマンスを維持できる。本システムの強化学習による環境変化への適応性を実証するため、学習の途中で空気圧制御系の電磁

弁を動作特性の異なるものに変更し、環境変化を引き起こす。このときのパフォーマンスの変化を調べた。

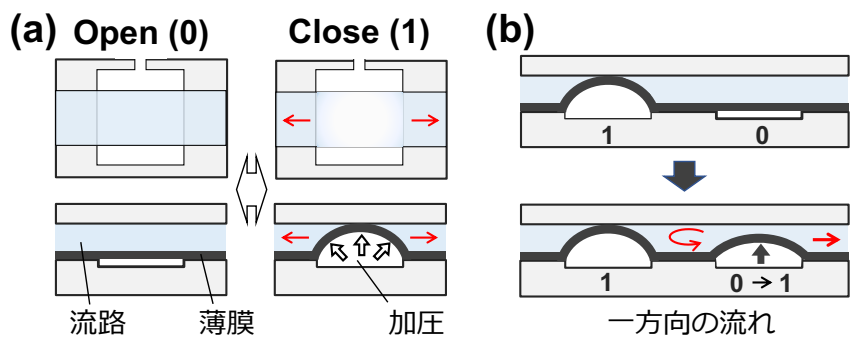


図 1.1 (a)マイクロバルブの構造 (b)マイクロバルブを用いた送液の原理

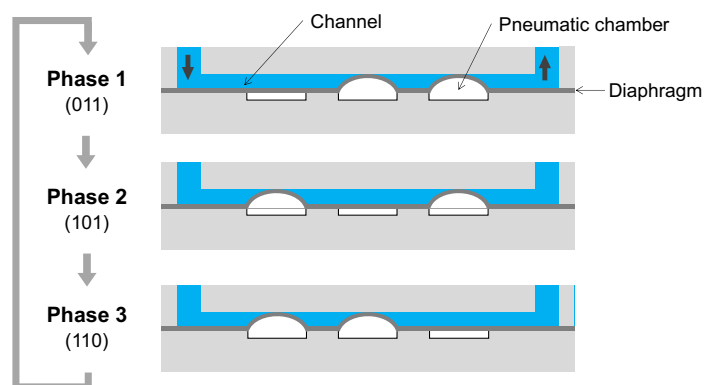


図 1.2 送液動作の例

1.2 理論、実験方法

1.2.1 デバイスの構造

図 1.3(a)にデバイスの構造を示す。本デバイスは、マイクロ流体システムにおけるマイクロポンプとして最小の構成である、3つのマイクロバルブで構成されるポンプを用いる。流路構造はこのポンプをリザーバを介してループ状とし、ポンプによって内部で液体が循環する。ポンプの右側に流量を観察するための領域を設け、図 1.3(b)のように内部のトレーサ粒子の移動を観察することで流量を測定する。操作領域の流路断面積を A 、粒子移動距離を x とすると、流量 V を以下のように定義する。

$$V \approx Ax \quad (1.1)$$

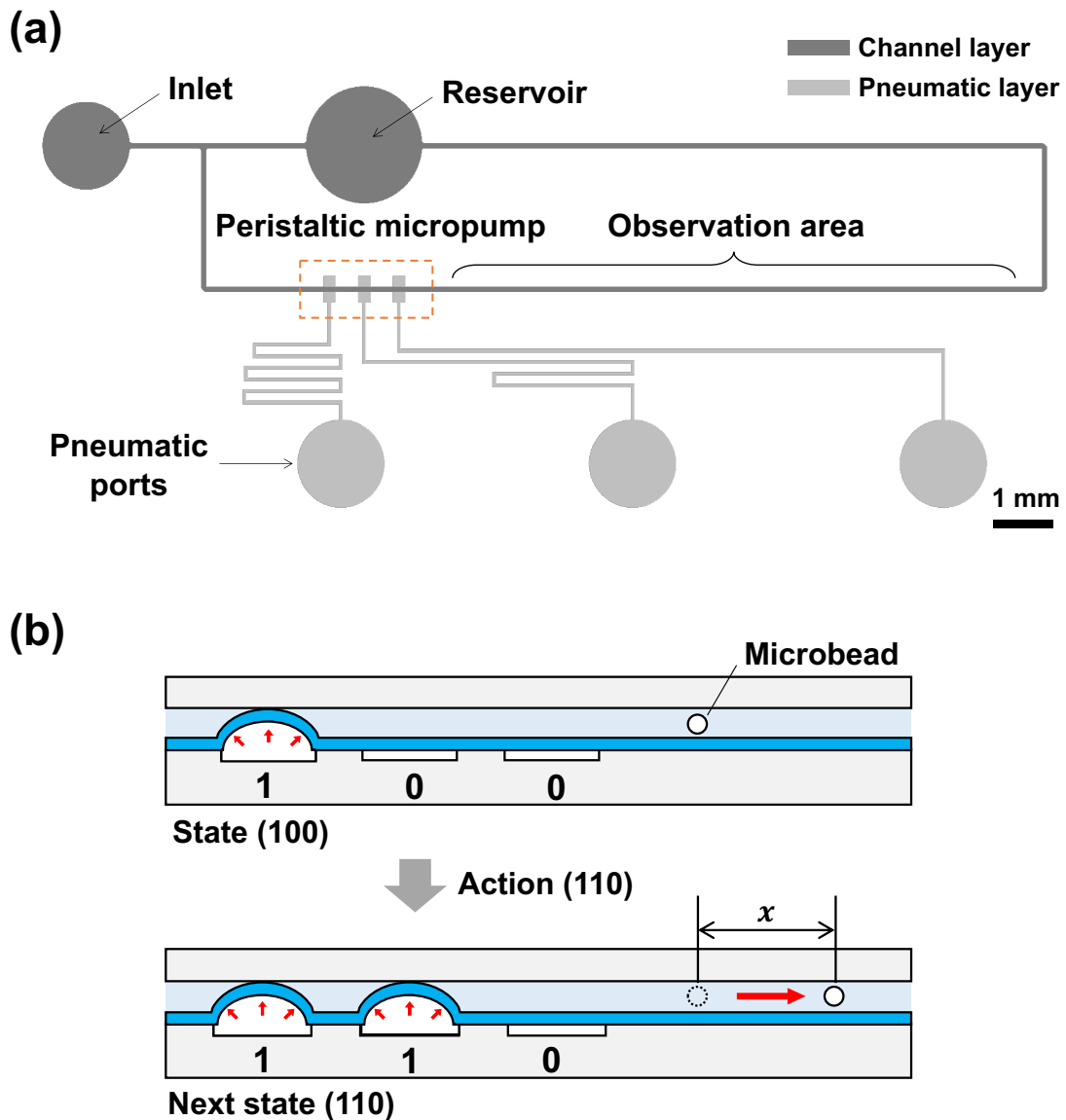


図 1.3(a) デバイスの流路構造 (b) 流量の測定

1.2.2 強化学習の適用

図 1.4 に強化学習の概要を示す。強化学習は、環境内で活動する行動主体(エージェント)を考え、エージェントがある状態 s で行動 a を実行した場合、環境から報酬 r が得られ、状態が s から s' に遷移する。その後エージェントの状態を $s = s'$ としてこれを繰り返す。エージェントの基本的な目的は、現在のタイムステップ t から終状態 ($t = T$) までに環境から得られる報酬の和

$$G_t = r_{t+1} + r_{t+2} + \dots + r_T \quad (1.2)$$

を最大化することである。ここで、タスクが無限に継続するものである場合、式(1.2)が無量大となる可能性がある。そこで、現在のタイムステップよりも未来で得られる報酬は、それが遠い未来であればあるほど不確実性があるため、価値を割り引いて考える。この考え方から、強化学習におけるエージェントは、割引された報酬和

$$G_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} \dots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \quad (1.3)$$

が最大となる行動を選択する。ここで、 γ は割引率と呼ばれる係数で、 $0 \leq \gamma \leq 1$ の範囲で設定される。

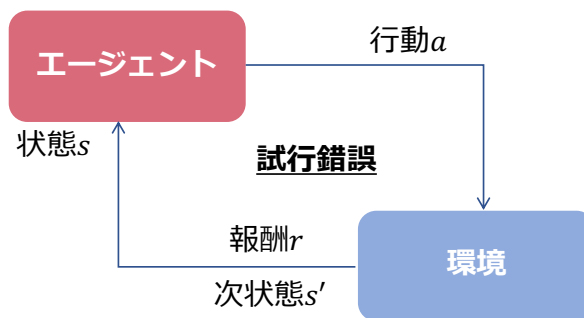


図 1.4 強化学習の概要

1.2.2.1 動的計画法

図 1.3(b)のように、マイクロバルブは Open/Close の 2 状態を取ることが可能で、それぞれ 0/1 として表記する。論理 1 (Close)は、空気室が加圧され、薄膜が流路を閉塞している状態を表す。一方、論理 0 (Open)は、チャンバ圧力が大気圧に等しく、薄膜は変形せず、流路が閉塞されていない状態を表す。

ポンプのダイアフラムの遷移によって発生する流量はダイアフラムの現在のフェーズと遷移後のフェーズのみで決まるとして、マルコフ性を仮定した。本実験系がマルコフ性を満たしている場合、MDP として定式化され、動的計画法に基づいて戦略を獲得することができる。MDP は、状態集合 \mathcal{S} 、行動集合 \mathcal{A} 、状態遷移確率関数 P 、報酬関数 R で定義される。マイクロペリスタルティックポンプにおける MDP の構成要素を以下のように定義する。

状態 $s \in \mathcal{S}$: ダイアフラムの現在フェーズ

行動 $a \in \mathcal{A}$: ダイアフラムの次フェーズ

報酬 $R(s, a)$: 状態 s で行動 a を実行した場合の流量

\mathcal{S}, \mathcal{A} はとり得るすべての状態、行動の集合である。

ここで、動的計画法で最適な行動価値を求めるためには、報酬関数が計算可能でなければならない。まず、ポンプの開閉状態は、空気圧制御系によって所望の状態へ遷移させることができるため、確率遷移しない。そのため、状態遷移確率関数 $P(s'|s, a)$ (s' は遷移後の状態を意味する) は考慮しない。上記の定義をもとに、Bellman 方程式から、状態 s における行動 a の価値 $Q(s, a)$ の式が得られる。

$$Q_{(s,a)} = R_{(s,a)} + \gamma \max_a Q_{(s',a')} \quad (1.4)$$

ここで、 γ は割引率である。

アルゴリズムは以下の通りである。

1. すべての状態 s と行動 a について、 $Q_{(s,a)}$ を 0 で初期化
2. すべての状態 s と行動 a について価値関数を更新。

$$Q'_{(s,a)} := R_{(s,a)} + \gamma \max_a Q_{(s',a')} \quad (1.5)$$

ただし、 $Q'_{(s,a)}$ は更新後の価値関数である。

3. 収束条件

$$\max_{s,a} |Q(s, a) - Q'_{(s,a)}| < \varepsilon \quad (1.6)$$

を満たしていれば、戦略

$$\pi_{(s)} = \arg \max_a Q_{(s,a)} \quad (1.7)$$

め、終了。それ以外は、 $Q(s, a) := Q'_{(s,a)}$ として 2 を繰り返す。

1.2.2.2 Q 学習の適用

環境に関する数理モデルが存在する場合、1.2.2.1 節で示したような動的計画法の適用により、最適な動作を得ることができる。しかし、現実の環境では、状態遷移や報酬関数は未知の場合が多い。また、動的計画法は状態行動空間のすべての価値関数を更新する必要があるため、空間が広すぎる場合に現実的な時間で計算することが難しくなる。このように、環境の数理モデルが得られない場合や状態行動空間が広大な場合は、有望な行動を中心に環境内の探索を行うアルゴリズムを用いる場合が多い。このような数理モデルを必要としない、モデルフリー型のアルゴリズムとして Q 学習[54]がある。この手法は、ある状態 s におけるある行動 a の価値 $Q(s, a)$ を、環境から得られる即時報酬 r から、以下の式で更新する。

$$Q(s, a) := Q(s, a) + \alpha \left\{ r + \gamma \max_{a' \in \mathcal{A}} Q(s', a') - Q(s, a) \right\} \quad (1.8)$$

ここで、 γ は割引率、 α は学習率である。本実験で用いられた Q 学習の計算手順は以下の通りである。

1. すべての状態 s と行動 a について、 $Q_{(s,a)}$ を任意の数値 Q_0 で初期化
2. 初期状態 s_0 を観測。
3. 方策モデル $\pi(a|s; Q)$ に従い行動 a を実行
4. 環境から報酬 r と次状態 s' を観測
5. 式(5) を用いて Q 値を更新
6. タスク終了条件到達で終了。それ以外は、 $s := s'$ として手順 3 から繰り返す。

ここで、エージェントの行動選択は方策モデルに従う。本実験では、基本的に現在の状態に対する Q 値が最大の行動を選択し、一定の確率 ε でランダムな行動が選ばれる。この手法は、 ε -greedy 方策と呼ばれる。

1.2.3 デバイスの作製

デバイスは、ソフトリソグラフィによって作製された[86]。まず、ネガ型フォトリソレジスト (SU-8 3025, Nippon Kayaku Co., Ltd., Japan) をシリコン基板にスピコートし、95 °C で 65 分加熱した。このとき、Channel layer の厚さは 28.5 μm であった。次に、このウェハをフォトマスクを使用して UV ライトに 250 mJ/cm^2 で露光した。その後、ウェハを 95 °C で 5 分間加熱し、現像した。バルブ作製のため、Channel layer のダイアフラム部分に厚さ 30 μm 、幅 100 μm のワックスを配置し、130 °C のホットプレートで 10 秒間加熱し、半円形の流路断面を得た。モールドの作製後、2 mm のシリコンゴムシートがスペーサとしてモールドに取り付けられ、デバイスの厚さを調整した。PDMS (SYLGARD 184, The Dow Chemical Co., USA) は、基剤と硬化剤を 10:1 の割合で混合し、モールドに流し込んだ。これを 75 °C で 90 分加熱することで硬化させた。モールドからはく離された PDMS は、外形をカットし、チャンバ一用の穴がパンチによって開けられた。ベントは 18G シリンジニードルを用いて開けた。ダイアフラムは 10:0.7 の割合で混合された PDMS をシリコンウェハ上にスピコートし、厚さを 23.7 μm とした。ダイアフラムと pneumatic layer に真空プラズマを照射し、貼り合わせたあとで 75 °C のホットプレートで 15 分加熱することで接合した。このレイヤーと channel layer を同じ手法で接合した。

1.2.4 学習システムの構築

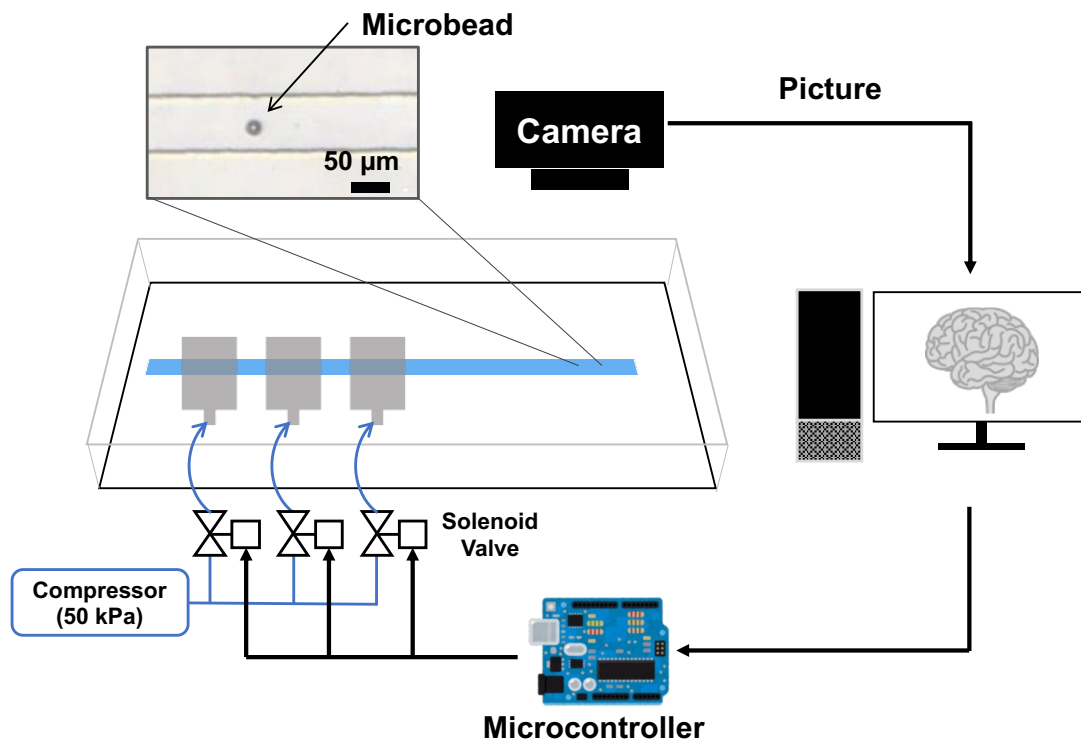


図 1.5 学習システム

学習システムを図 1.5 に示す。システムは、マイクロ流体デバイス、ダイヤフラムを駆動するためのソレノイドバルブ(USG3-6-1-DC12V, CKD Corporation, Japan)、ソレノイドバルブを制御するための Arduino Microcontroller (Arduino Uno, Arduino S.R.L., Italy)、学習処理を行うコンピュータ、顕微鏡と CMOS カメラ(VC-320, Gazo Co., Ltd.)で構成される。Pneumatic chamber に印加される圧力は、本章のすべての実験において 50 kPa とした。デバイスの流路内を、Milli-Q water に Tween20 (170-6531, Bio-Laboratories, Inc., USA) を 0.05%(v/v)で希釈したもので満たした。その後、デバイスの流路 20 μm のポリスチレン粒子(#18329, Polysciences Inc., USA)を入れた。CMOS カメラから画像をコンピュータに取り込み、OpenCV 4.0 を用いて粒子を抽出し、粒子位置を取得する。

1.2.5 流量のサンプリング

流量サンプリング方法は次のとおりである。まず、ダイヤフラムを(000)に遷移させ、状態をリセットする。その後、ダイヤフラムを状態 s に遷移させ、粒子位置を測定する。その後、状態を s' に遷移させ、同じように粒子位置を測定する。ここで、遷移前後で測定した粒子位置の差と流路の断面積から、移動した流体の体積を算出した。

上記の測定をすべての状態 s と行動 a の組み合わせで測定することで、報酬関数 $R(s, a)$ として利用した。また、1回の遷移あたりの動作速度は、バルブの開閉速度よりも十分に遅いため、粒子の移動量の合計を流量に近似できると仮定した。

サンプリングした流体の移動量から、シーケンスの流量を推定する手法として、状態行動の組に対する流量を合計する。図 1.6 に流量推定の例を示す。初期状態 s_0 で行動 a_1 を実行した場合に報酬(流量) r_{01} が生じる。この行動により状態が s_1 に遷移する。3 フェーズのシーケンス(s_0, s_1, s_2)の場合、終状態が初期状態 s_0 に一致するときの合計流量($r_{01} + r_{12} + r_{20}$)をそのシーケンスの流量とする。これを遷移回数で除したものがそのシーケンスの1ステップあたりの平均流量となる。例えば、(011)(101)(110)の3 フェーズシーケンスの場合、(011)→(101), (101)→(110), (110)→(011)の3 遷移で発生する流量を合計することで1 循環あたりの発生流量を求めることができる。

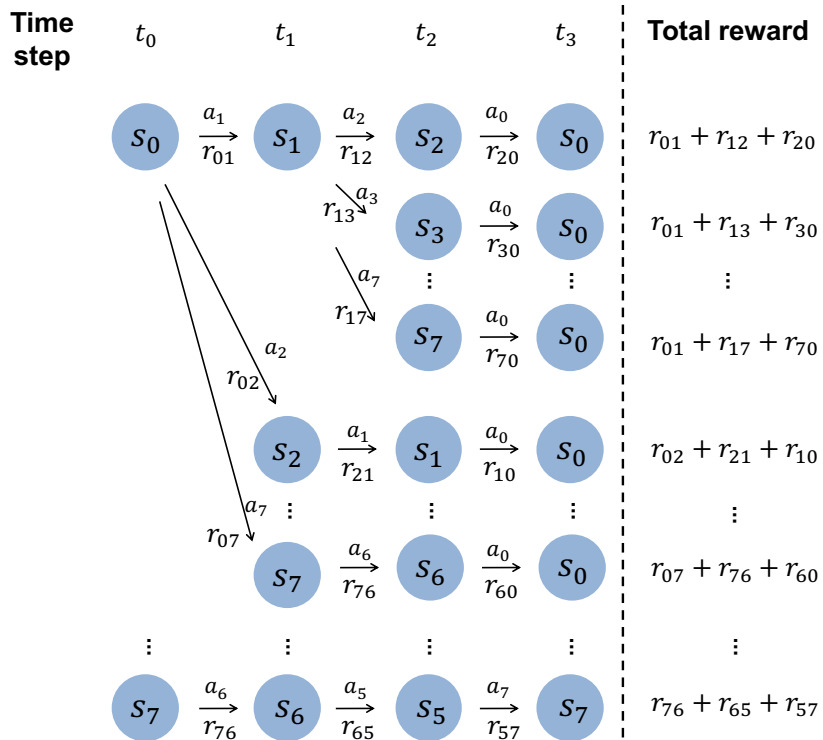


図 1.6 シーケンスの流量推定方法

1.3 実験結果および考察

1.3.1 動的計画法による動作シーケンスの獲得

動的計画法の適用のため、報酬関数を決定する必要がある。マイクロバルブは現在のバルブ開閉状態から次の開閉状態に遷移するとき流体の移動が発生する。そのため、状態行動のすべての組み合わせに対して移動する流体の量を測定することで、報酬関数を定義した。表 1.1 に測定された粒子移動量を示す。

サンプリングした流量を報酬関数として用いて動的計画法により最適な行動を求めた。図 1.7(a)は割引率を変えた場合の残差の推移を示している。割引率が 1 に近づくと、より遠いタイムステップでの報酬が価値評価に影響を及ぼすため、一つの価値を決定するためのファクターが増加し、残差の減少が緩やかになる。マイクロポンプの動作は、1 回の動作で発生する流量よりも複数回の動作で発生する流量の合計値が重要であるため、割引率は 1 に近いほうが理想的である。しかし、割引率が上昇すると集束までの時間が増加するため、適切な調整が必要である。図 1.7(a)より、割引率 $\gamma = 0.70 \sim 0.90$ では、100 iteration 以内に収束した。 $\gamma = 0.99$ の場合、約 700 iteration で収束条件に達した。図 1.7(b)は $\gamma = 0.80, 0.90$ の場合の政策収束の経過を示す。数回の反復で戦略が収束していることがわかる。図中の矢印は、戦略によって決まる遷移の方向を示している。図 1.8(b)は各割引率設定で収束した最適な行動を示している。 $\gamma = 0.70, 0.80$ と $\gamma = 0.90, 0.99$ はそれぞれ同じ戦略となることが確認された。図 1.9 は従来提案されていた動作シーケンスと動的計画法により得られたシーケンスの流量の比較である。図にはサンプリングされた流量から計算された値（シミュレーションのプロット）を同時に示した。

強化学習によって得られた(000)(110)(001)のシーケンスでは、従来提案されている 4 または 6 フェーズのシーケンス[87]に比べ、流量が 2 倍に増加した。 $\gamma = 0.70, 0.80$ で収束した(110)(101)は左のダイヤフラムが常に閉じているため、流れが発生しないことが示された。また、従来の 3 相配列 (011)(101)(110)のシミュレーション値と実測値は大きな不整合を示している。これらの結果は、以下で説明するように、ダイヤフラムの開閉動作の速度差によるものである。

図 1.10 は、バルブの動作の瞬間を撮影した画像である。制御信号が入力された後、中央のダイヤフラムは変化が確認されてから約 15 ms で閉じるが、右側のダイヤフラムは 30 ms で開き始め、45 ms で完全に開放された。この動作速度の差は、ダイヤフラムの空気圧を制御する電磁弁の応答速度に起因するものである。また、ダイヤフラムの開放状態への変化は PDMS の弾性にのみ依存することも影響していると考えられる[85]。

2 フェーズのシーケンス (110)(101) における(110)から(101)への遷移では、 $1 \rightarrow 0$ が $0 \rightarrow 1$ よりも遅いため、実質的に、(110)→{(111)}→(101)となる（{} は遷移途中の物理的なバルブ状態を示す）。この遷移単体で見れば、右のバルブに保持された流体が右に流出することとなり、流れが生じる。これによってシミュレーション上では大きな流量が生じると算出されるが、実際には左のバルブは閉じているため、シーケンス全体では流れは生じない。

従来の3相シーケンス(011)(101)(110)の流量は、他のシーケンスと比較して極端に少ない。この事実は、シミュレーションとの間に大きな矛盾がある。この矛盾は、次のようなプロセスで生じる。まず、(011)(101)の遷移時に、ダイヤフラムの開閉速度の違いにより物理遷移が(011)→{(111)}→(101)となり、中央のダイヤフラムは実質的に常に閉じた状態が維持される。その後、(101)→(110)において、{(111)}→(110)となる。このとき、中央のバルブには流体がほとんど保持されていないため、右の流路からの逆流が生じた。このように、過去の動作履歴による制御信号と物理状態の不整合により、従来の3相シーケンスでは、シミュレーションと一致する流量が得られなかったと推察される。そのため、図1.9における、ダイヤフラムの動作遅れの影響を受けないと考えられる3フェーズシーケンスでは、シミュレーション値と実測値がよく一致した。ここで、シミュレーションにおいて、すべての3フェーズシーケンスの流量を算出した場合、本実験で獲得したシーケンスが最も流量の多いものであることがわかった。

これらの結果から、マイクロペリスタルティックポンプの運転にマルコフ特性を仮定し、MDPに基づいて定式化することにより、ポンプの運転に強化学習を適用することができることが示された。

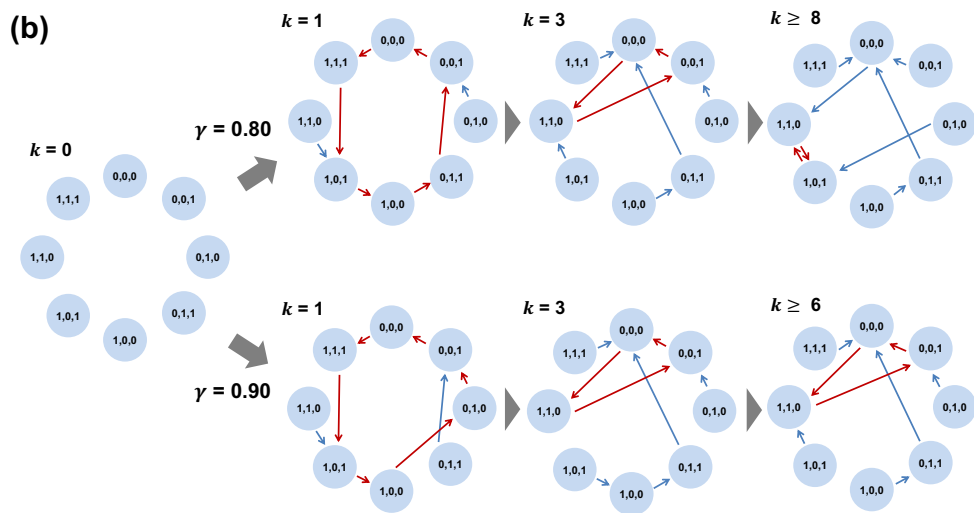
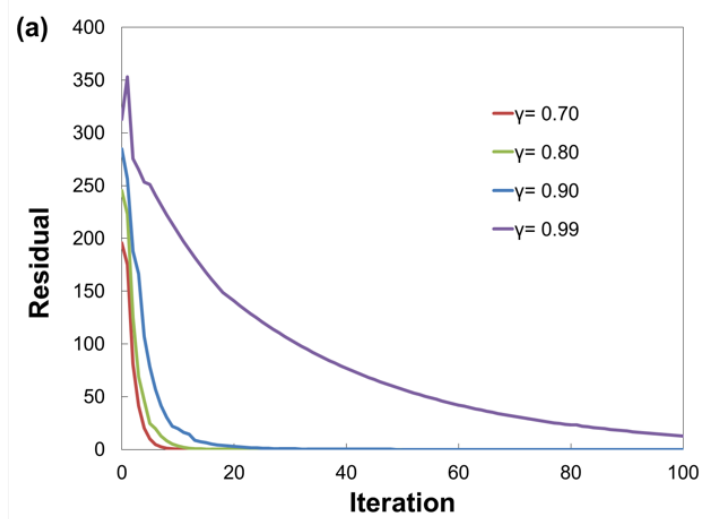


図 1.7 (a) 試行回数に対する最大残差の推移 (b)各割引率の設定で収束した戦略。図中の 8 つの円はポンプの状態を示し、矢印は戦略関数により与えられた行動 (次に遷移すべき状態) を示している。任意の初期状態から戦略に従い遷移すると 2 及び 3 つの状態を循環する (赤い矢印)。どの状態を初期状態としても戦略に従って遷移を繰り返すことで、(110)(101)か (000)(110)(001)の循環に収束する (赤い矢印)。

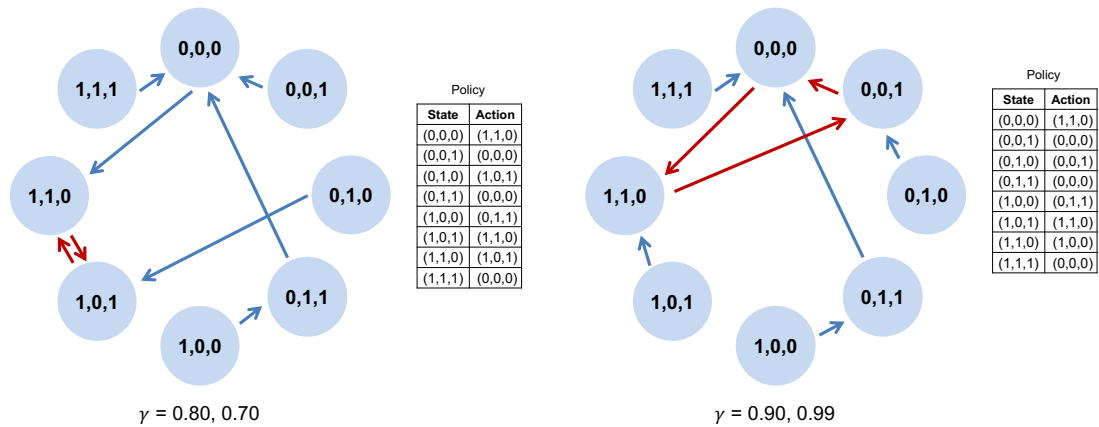


図 1.8 各割引率で収束したシーケンス

表 1.1 各状態行動における粒子移動量

			Action (next state)							
			0	1	2	3	4	5	6	7
			(0,0,0)	(0,0,1)	(0,1,0)	(0,1,1)	(1,0,0)	(1,0,1)	(1,1,0)	(1,1,1)
State	0	(0,0,0)	2.1	91.2	75.5	139.9	109.9	157.4	216.3	249.1
	1	(0,0,1)	-57.1	-1.0	-225.9	-3.4	-180.6	-3.4	-136.6	-5.4
	2	(0,1,0)	-40.2	204.4	-3.1	202.7	-155.2	203.0	8.8	202.7
	3	(0,1,1)	-88.7	2.0	-239.2	-0.6	-408.6	-0.8	-224.7	-3.3
	4	(1,0,0)	-45.4	165.1	168.7	339.2	-2.1	161.8	181.5	346.1
	5	(1,0,1)	-83.7	2.0	-169.6	0.1	-164.1	-0.3	24.8	96.0
	6	(1,1,0)	-100.0	208.2	2.7	206.2	-177.9	207.6	-0.4	204.0
	7	(1,1,1)	-117.0	4.2	-188.4	1.7	-350.8	1.7	-164.9	-0.7

Unit: [μm]

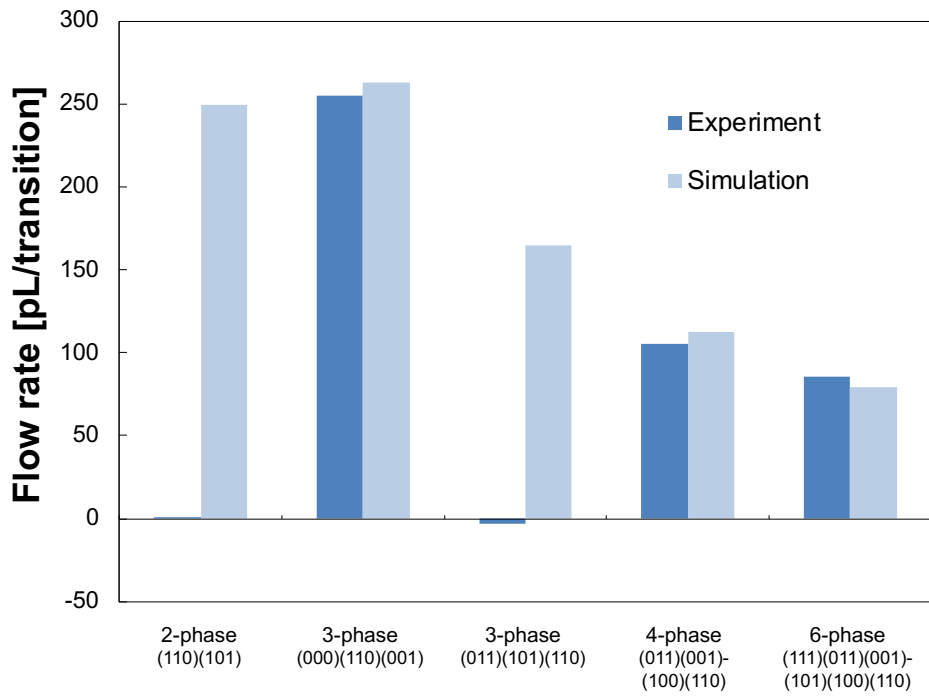


図 1.9 各動作シーケンスの流量の比較

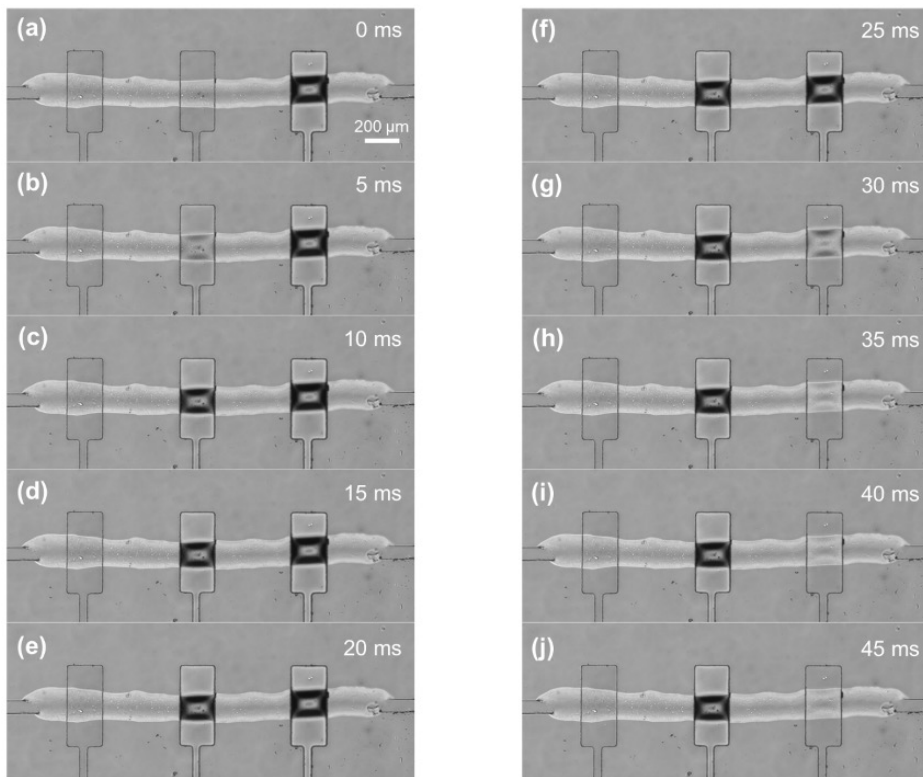


図 1.10 ダイアフラムの動作の連続写真。(001)から(010)への遷移の様子。

1.3.2 Q 学習による動作シーケンスの獲得

前節で、MDP の適用が可能であることが示された。ここでは、報酬関数が未知であることに加え、確率的に変化する場合に対応するため、環境を探索することで学習を行うアルゴリズムである、Q 学習の適用を試みた。

図 1.11(a)は、学習過程におけるダイアフラムの 1 遷移あたりに発生する流量の推移である。学習初期はほとんどランダムな遷移であるため負の流量（逆流）が確認されている。その後の学習によって、発生する流量は 1 step あたり約 216 pl となった。約 10000 step の学習で得られた動作シーケンスを図 1.11(b)に示すどの状態を初期状態としても戦略に従って遷移を繰り返すことで、(000)(110)(001)の循環に収束する（赤い矢印）。これは、2.3.1 節で動的計画法を用いて獲得した最適な動作に一致する。この流量の増加は、使用したデバイスの閉じた流路レイアウトやバルブの空気圧を制御する電磁弁の応答速度に起因するものであり、本実験系の特性が考慮された動作であることが確認された。

Q 学習の適用により、目的に沿ったシステムの動作を環境内の探索により最適化できることが示された。また、環境との相互作用によって個々のシステムに最適化された動作を得られることも示された。この技術は、プラットフォームのロバストな制御や、商業化における低コスト化、ユーザビリティの向上に大きく貢献する技術である。

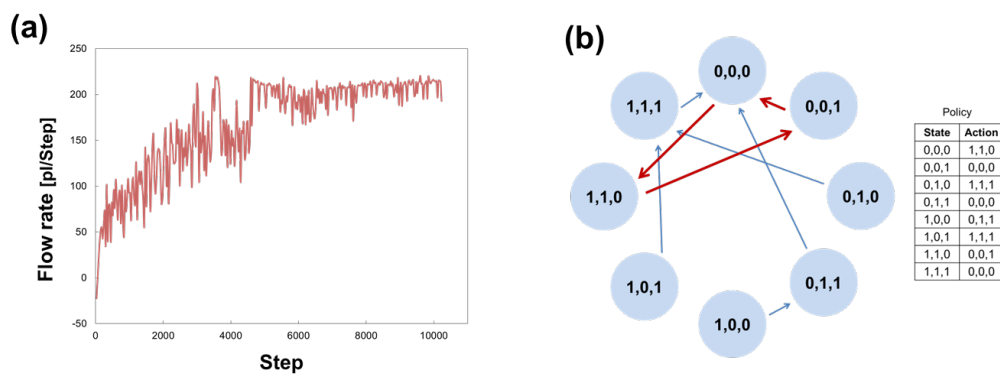


図 1.11 (a)1 ステップ当たりの流量の推移。 $\gamma=0.9$, $\alpha=0.1$ 。グラフは $N=30$ の移動平均。
(b) 獲得された戦略。丸はポンプのダイアフラムの状態、矢印は次に遷移すべき状態。

1.3.3 Q 学習による流量調整動作の獲得

エージェントはダイアフラムの現在の状態から行動価値関数が最大となるような行動を出力し、ダイアフラムが遷移する。流量に応じた報酬がエージェントに与えられ、Q 値の更新式に従って行動価値が更新される。この報酬の設定によって、エージェントの目的は粒子移動量（流量）を目標範囲に保持するような行動を学習することと定義される。

図 1.12 (a) は各流量設定における学習ステップ数と 1 ステップ当たりの流量推移である。学習初期における行動は、行動価値の学習がほとんど行われていないため、ランダムに近い遷移をとる。そのため、粒子の逆方向の移動（負の流量）が観測される。その後、学習を重ねるにつれて粒子の順方向への移動量が増加し、指定された流量範囲付近に流量を維持できることが確認された。ランダムな遷移では順方向への有意な流れを継続的に生み出すことはできないため、タスクスタートから順方向への流量を増加させる必要がある。そのため、目標流量が小さい場合では、0 から 10~20 pL/step のグリッドまでを探索すれば良いが、目標流量が 70~80 pL/step の場合は、0 からその間の状態を探索する必要があるため、より学習に時間がかかる。

学習終盤においても流量は目標範囲に対し最大 20 pL/step 程度のばらつきがみられた。今回のタスク設定では、報酬は過去の遷移に影響を受けるため、過去の行動によって現在の状態に関わらず最適な行動は異なる場合がある。平均流量の算出範囲は 50 遷移であるため、1 step あたりの影響は少ないが、過去の履歴によって最適な行動が変化し、一時的に流量が目標から外れ、これを修正するという行動が繰り返されることがばらつきの原因と考えられる。図 1.12 (b) は、それぞれの目標設定において、現在の状態が目標流量範囲内の場合のエージェントが獲得した戦略である。図中の矢印は現在のバルブ開閉状態において最も Q 値の高い行動を示しており、基本的にエージェントはこの矢印に沿って行動する。現在の状態がどのバルブ状態であったとしても、戦略にしたがって行動することで一定の繰り返し遷移（赤い矢印）に収束する。図より、各目標値によってそれぞれ異なるシーケンスに収束しているため、目的に応じた戦略を学習しているといえる。ここで、peristaltic pump では動作速度の変更によって流量を調整することが一般的である。しかし、本実験で示された、目的に応じた独自の動作パターンの獲得は、これまでにないマイクロ流体素子の応用方法を提示できる可能性を示唆している。

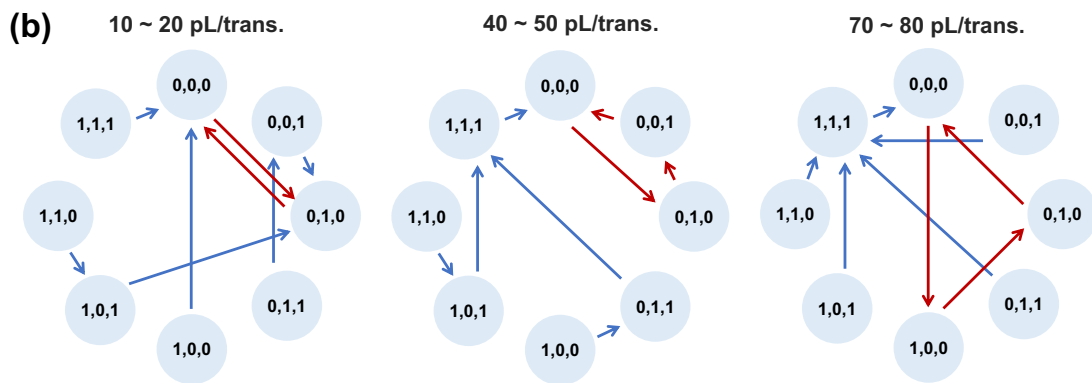
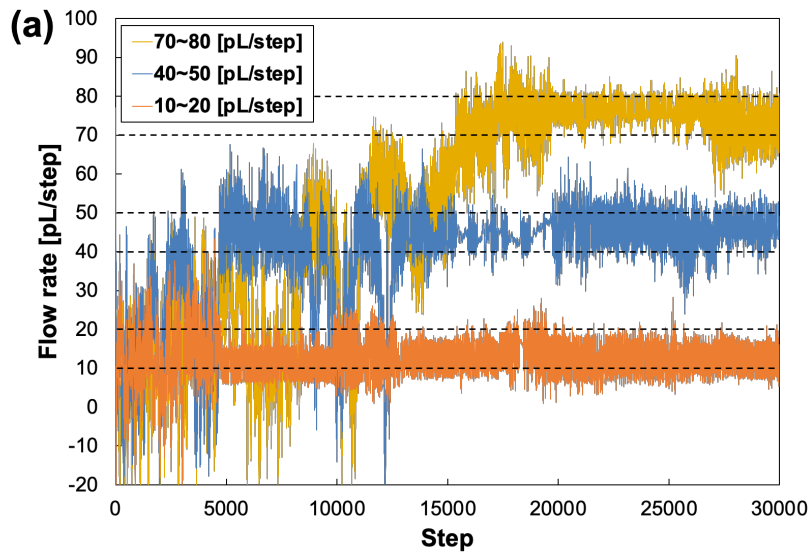


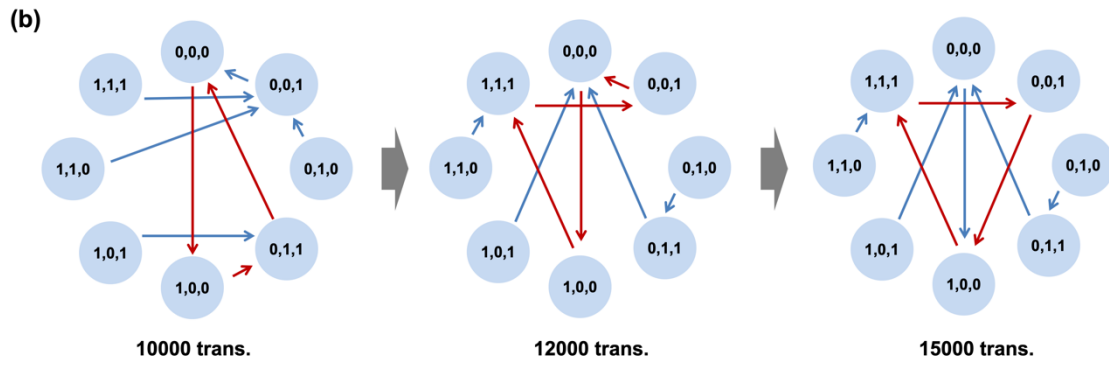
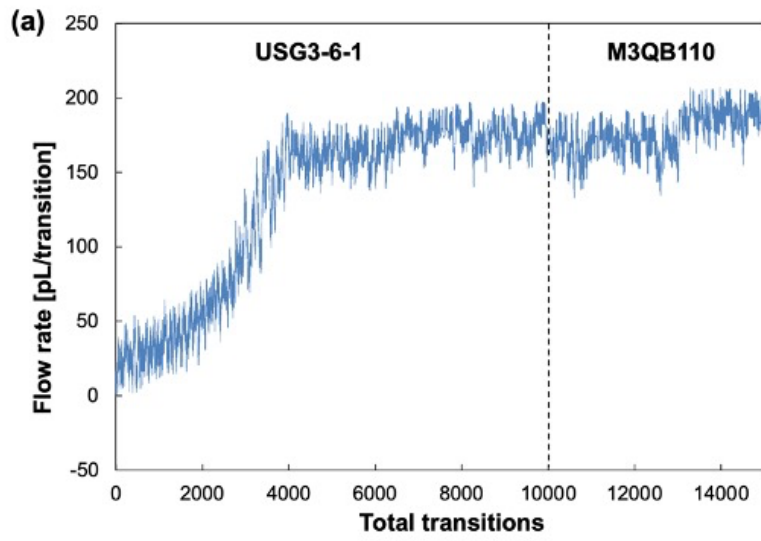
図 1.12 (a) 学習中の流量の推移($\gamma = 0.9, \alpha = 0.3, \varepsilon = 0.01$)。破線はそれぞれの目標流量範囲を表す。(b) 30000 step の学習後の戦略。円はポンプの状態を示し、矢印は戦略関数により与えられた行動(次に遷移すべき状態)を示している。任意の初期状態から戦略に従い遷移すると2及び3つの状態を循環する(赤い矢印)。

1.3.4 強化学習による環境変化への適応性の実証

強化学習は、環境内を探索し、価値観数を更新し続けることで未知の環境での最適な行動を獲得する。この探索によって、環境が変化した場合でもパフォーマンスが維持できるような行動を選択し、その環境に最適化された価値評価を再度学習できる。これは、ロバストなマイクロ流体システムの実現に寄与する。これを実証するため、ポンピングの学習の途中で、バルブ制御に用いられるソレノイドバルブを動作特性の異なるものに変更し、環境変化を引き起こした場合のパフォーマンスを調べた。

ソレノイドバルブは、学習開始から 10000 step までは USG3-6-1 (CKD, Japan) が使用され、それ以降は M3QB110 (CKD, Japan) が用いられた。図 1.13(a) より、学習開始から平均流量が増加し、4000 step ほどで横ばいとなった。その後、学習が 10000 step に到達した段階でソレノイドバルブを交換し、学習が再開された。図 1.13(b) は、学習中の戦略の推移を示している。エージェントの戦略は、10000 step の学習で(000)(100)(011)であった。ここで、これまでの実験における最適なシーケンスと収束結果が異なるのは、デバイスの個体差や実験環境の影響と考えられる。ただし、このシーケンスの送液のメカニズムは本文のシーケンスと同様にソレノイドバルブの開閉速度差を利用しているため、環境に応じたシーケンスが選択されていると考えられる。

図 1.14 に各シーケンスのそれぞれのソレノイドバルブにおける 1 遷移あたりの流量を示す。交換後のバルブを用いると、前のソレノイドバルブで獲得したシーケンスでは流量が大幅に低下した。これは、ソレノイドバルブを変更することが、環境変化を引き起こすことを示している。しかし、エージェントは戦略を切り替えることによってパフォーマンスを維持することに成功している。15000 step の学習後、エージェントは(001)(100)(111)の戦略をとっていた。状態動作のすべての組み合わせについて粒子の動きをサンプリングして流量を推定したところ、M3QB110 を使用した場合、このシーケンスが最大の流量を生み出すことを確認した。したがって、強化学習を用いることで、システムが環境変化に適応可能であることが実証された。



☒ 1.13 (a) Variation of the flow rate per transition. $\gamma = 0.9$, $\alpha = 0.3$. (Moving average, 50 trans.)
 (b) Operation sequence given by $\pi(s) = \arg \max_a Q_{(s,a)}$. The numbers under the figures indicate the total transitions.

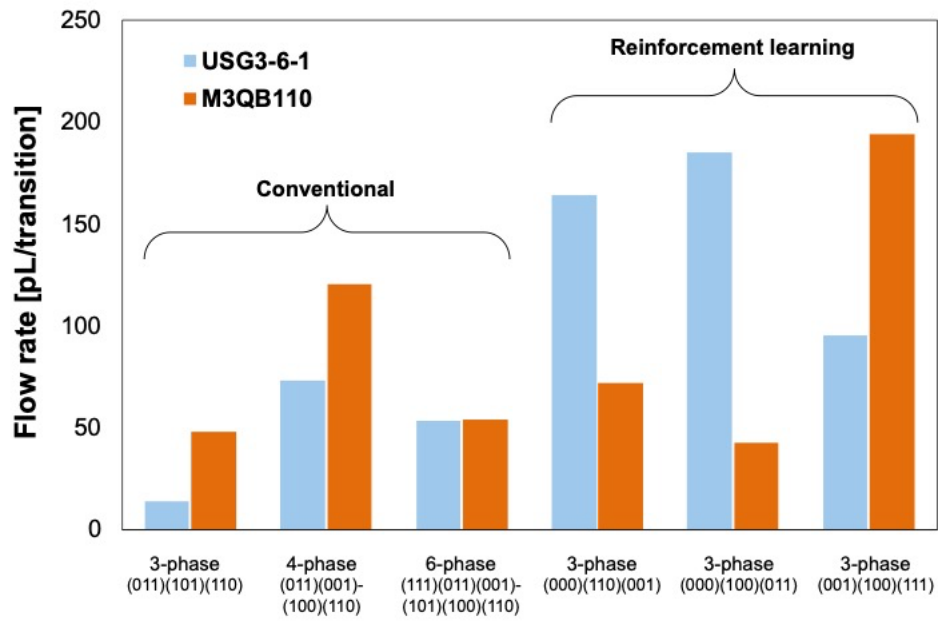


Fig. 1.14 Comparison of flow rate with each actuation sequence. These flow rates indicate the averaged flow rate per single transition, which is calculated from the measured total flow rate with a transition frequency of 3 Hz.

1.4 結言

本章では、オンチップマイクロバルブの制御に対する強化学習の適用を試みた。制御対象として、3つのマイクロバルブを搭載したマイクロペリスタルティックポンプを用いた。

まず、マイクロペリスタルティックポンプをマイクロ流体デバイス上に構成し、このデバイスに周辺機器を接続し、学習プラットフォームを構築した。強化学習を適用するため、マイクロポンプの動作にマルコフ性を仮定し、MDPとして定式化した。動的計画法により、最適な動作を決定した。バルブの動作で発生する順方向への流量をMDPにおける報酬と定義することで、エージェントの目的を流量の最大化に設定した。動的計画法の適用のため、ポンプの動作で発生する流量をすべての動作パターンでサンプリングすることにより、報酬関数がモデリングされた。これにより、コンピュータ上で学習が行われた。割引率の設定によって、収束する動作がことなることが示された。割引率を適切に設定することで、(000)(110)(001)の3フェーズシーケンスが得られ、従来のシーケンスに比べて50%以上流量が増加した。加えて、使用したデバイスや周辺機器の動作特性に応じた行動の選択が行われていることを確認した。

その後、環境が未知の場合、報酬関数が定義できない場合、モデルフリーの学習手法を適用することで環境を探索し、最適な行動を学習することができる。代表的なアルゴリズムとして、Q学習の適用が行われた。サンプリングされた流量をもとに作成した報酬関数を用いてコンピュータ上での学習と、実環境での学習を行い、探索による送液動作の獲得が確認され、動的計画法で獲得されたものと同様のシーケンスに収束した。

より高度な制御として、ポンプの流量を所望の値に調整する動作の獲得を示した。報酬を目標流量からの偏差を基準に算出することで、所望の平均流量を発生する動作を獲得できることを示した。

さらに、Q学習の探索による環境変化への適応能力を示すため、学習の途中で空気圧制御系の電磁弁を変更し環境を変化させた。エージェントは戦略を切り替えることにより、環境変化に対してパフォーマンスを維持できることが示された。環境の変化後に獲得されたシーケンスは、シミュレーション上、最大の流量を生み出すことを確認した。したがって強化学習による環境変化への適応能力が示された。

第2章 強化学習を利用した1次元微粒子操作技術の開発

2.1 緒言

前章では、オンチップマイクロバルブの制御に対する強化学習の適用を行い、動作の自律的学習と効率化を実証した。

本章では、より高度なタスクとして、同一のシステムによる粒子操作技術への拡張を試みた。マイクロバルブは動作の組み合わせによって流体の流れが生じる機構である。流路内に操作対象の粒子を配置し、バルブを適切に動作させることで、マイクロバルブを利用した粒子操作技術が実現できる。どのようにバルブを動作させれば粒子を目的の位置に配置できるかは未知であるが、強化学習における報酬の設定を変更すれば、自律的に獲得が可能である。

ここで、Q学習などの探索を必要とするアルゴリズムでは、学習は試行錯誤によって行われるため、タスクの難易度が上昇すると膨大な数の試行錯誤が必要となる場合がある。そのため、実環境における学習では、現実的な時間で学習が終了しない可能性がある。よって、本章においてもシミュレータを構築し、学習の効率化の検討を行った。実環境でゼロから学習したモデルとシミュレータで事前学習されたモデルを比較し、強化学習を用いた新しい粒子操作技術の実証と、大規模化を視野に入れた学習環境の評価が行われた。

2.2 理論、実験方法

2.2.1 強化学習の適用

マイクロポンプの強化学習における状態と行動は以下の通りである。

状態 $s \in \mathcal{S}$: ダイアフラムの現在フェーズ、粒子位置

行動 $a \in \mathcal{A}$: ダイアフラムの次フェーズ

\mathcal{S}, \mathcal{A} はとり得るすべての状態、行動の集合である。

Q 値の定義は式(5)と同様である。

本実験で用いられたアルゴリズムは以下の通りである。

1. すべての状態 s と行動 a について、 $Q_{(s,a)}$ を任意の数値 Q_0 で初期化
2. 初期状態 s_0 を観測。
3. 方策モデル $\pi(a|s; Q)$ に従い行動 a を実行
4. 環境から報酬 r と次状態 s' を観測
5. 式(5) を用いて Q 値を更新
6. タスク終了条件到達で終了。それ以外は、 $s := s'$ として手順 3 から繰り返す。

ここで、方策モデル ϵ -greedy 法を使用し、 $\epsilon=0.05$ とした。

2.2.2 粒子操作シミュレーションおよび実験環境

粒子操作タスクの定義を図 2.1 に示す。1 次元の直線流路を仮定し、粒子の中心位置をターゲット領域に一致させることを目的とする。粒子は中心位置のみが考慮される。強化学習の適用のため、粒子の位置が状態表現に組み込む。本章で用いる Q 学習は状態を離散化する必要があるため、流路にグリッドを設定し、粒子の位置をグリッド番号で表す。まず、ターゲットの位置を基準 (Grid 0) として、ターゲットの両側にグリッドを定義する。ここで、流路内の流れは流路の x 方向位置に関わらず一定と仮定し、これに従い、粒子の移動はバルブにより生じる流体の移動量にのみ依存し、これは流路の x 方向位置に依存しないとする。最終的なタスクは流路の x 方向の任意の位置に粒子を配置することであるが、シミュレータ上では実質的に、粒子を初期位置から Grid 0 に移動するタスクと定義できる。このタスクをシミュレータ上で学習できれば、実環境に置いて任意の位置への操作が可能となる。

エージェントが得る報酬 r は、以下の通りである。

$$r = \begin{cases} 1 & (s_p = 0) \\ -1 & (\text{otherwise}) \end{cases} \quad (2.1)$$

グリッドの定義範囲外に粒子が移動した場合は、ポンプの状態は遷移せず、最も外側のリセットエリアに戻されるとし、この場合も -1 の報酬が与えられる。シミュレーションのための粒子移動量のサンプリングは 1.2.5 と同様である。

実環境における粒子操作の実験では、設定された範囲内のランダムな位置にターゲットを定義し、タスクを実行する。初期バルブ開閉状態もランダムに決定される。学習中のシステム動作速度は 2 Hz とした。ターゲットの左右に片側 10 グリッドが定義され、グリッドの幅はカメラ解像度を基準に 20 px (45.8 μm) とした。実験中に粒子がグリッド定義範囲から外れた場合のため、この 10 グリッドの外側に幅 100 px (229 μm) のリセットエリアが設定された。粒子は固定されたシーケンスを用いてこのエリア内のランダムな位置に戻される。よって、定義されるグリッド数は、ターゲット、リセットエリアを含めて 23 グリッドとなる。

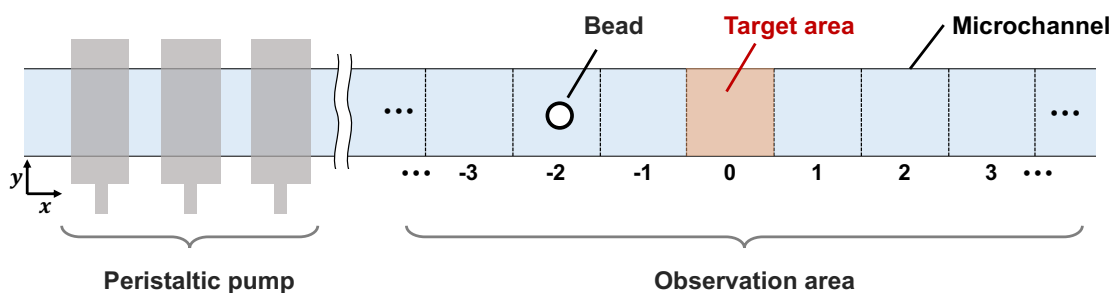


図 2.1 粒子操作タスクの定義。グリッド幅はカメラ解像度で 20 pixel (45.8 μm)。

2.3 実験結果および考察

まず、コンピュータ上で学習を行うため、シミュレータを構築した。シミュレータは、サンプリングした微粒子の移動データをもとに、微粒子の移動を予測する。学習の進行は、初期位置から目標に到達するまでに要した動作回数（ステップ数）を用いて可視化した。

図 2.3 のプロットは、シミュレーションと実世界の学習結果を示している。各 episode はエージェントが目的を達成するまで続けられた。実験のコントロールとして、エージェントの行動をランダムにした場合の結果も示している。これは、同じタスクにおいて、すべての行動をランダムに選択した場合の結果である。シミュレーションでは、学習が進むにつれて目標に到達するまでのステップ数が減少し、評価に用いた 100 episode の平均値は、合計 1000 episode の学習で 3.0 steps/episode に収束した。しかし、ランダムアクションの場合、45 episode に到達するまでに約 6,000 steps を要し、平均値は 135.9 steps/episode となった。この結果は、エージェントがシミュレータでタスクをクリアするために効率的な行動を学習したことを示している。

図 2.4 は、実験中の流路とポンプのフェーズの step-by-step の画像である。事前学習を行わない場合、最初の 100 episode の平均値は 28.5 steps/episode であったが、約 1000 episode の学習後、最後の 100 episode の平均値は 4.6 steps/episode であった。しかし、事前学習を行った場合、最初の 100 episode の平均値は 9.0 steps/episode となり、事前学習を行わない場合よりも効率的であった。さらに、1,000 episode に到達するまでの総 steps 数は、2,500 step 以上削減された。これは、エージェントの学習コストの削減とシステム規模の拡大に寄与している。ここで、2つのエージェントのタスク性能は、シミュレーションで到達した最終値と比較して低下しているように見える。実環境では、温度などの環境条件の変化や、流路の y 方向の位置に関する微粒子の移動などを考慮したポリシーが必要である。しかし、シミュレーションではこれらに起因する摂動が実装されていないため、結果にギャップが生じたと考えられる。

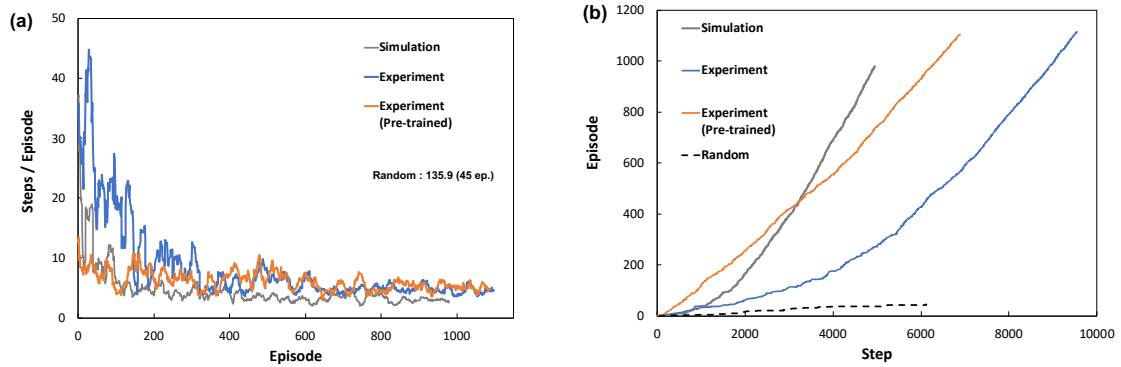


図 2.2 (a) 1 回のタスククリアに要した行動回数の推移($\gamma = 0.9$, $\alpha = 0.1$, $Q_0 = 0$)。プロットは $N=20$ の移動平均を示す。(b) トータルステップ数に対するエピソード数。

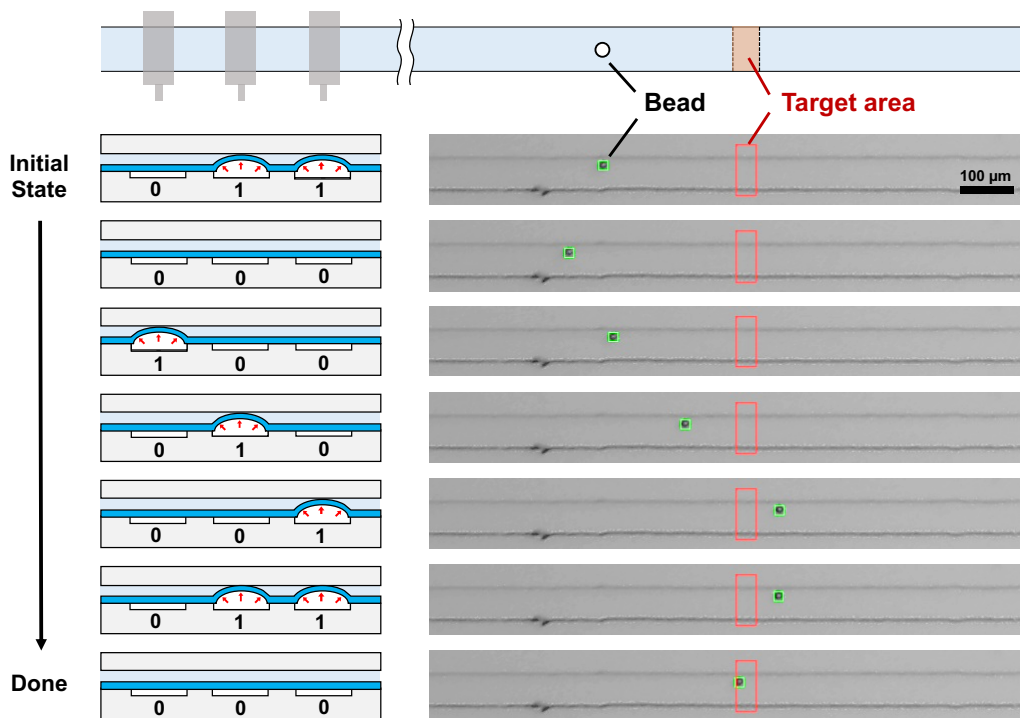


図 2.3 粒子操作タスク実行中の連続写真とポンプの動作

2.4 結言

本章では、前章までに述べた、マイクロ流体システムに対する強化学習の適用手法を拡張することで粒子操作技術を実証した。マイクロ流体システムに強化学習を適用することで、特微量を変化させるだけで、環境との相互作用から与えられたタスクを実行するための制御方法を自律的に獲得することが可能となる。

強化学習を適用するため、特微量とタスクの設定が行われた。本粒子操作タスクは、直線状の流路に配置された微粒子を指定されたターゲットエリアに運ぶタスクとした。粒子がターゲットエリアに到達した場合に正の報酬を与えることで、粒子を目的位置に運ぶためのバルブ開閉制御の学習を目的とした。

実環境での適用に先立ち、コンピュータ上での学習を行うためのシミュレータを構築した。これは、サンプリングされた粒子移動量を用いてコンピュータ上でタスクを再現するものである。学習アルゴリズムには Q 学習が用いられ、シミュレータ、実環境のそれぞれの環境で学習が行われ、粒子を効率的に運ぶ動作を獲得していることが示された。また、シミュレータによる事前学習によって実環境における学習時間を削減可能であることが示された。

本技術は、流路内のウェルや堰などの物理的構造を必要とせず、高価な機器を使用しないため、低コストで自由度の高い技術である。今後、粒子操作領域を 2 次元平面に拡張し複数粒子の並列操作を実現することで、低コストでハイスループットな粒子配列技術としての応用が期待できる。

第3章 Vacuum-assisted Micromolding を用いたリフロープロセス

3.1 緒言

前章までに使用したマイクロバルブは、マイクロ流体デバイスにおける流体制御を担う重要な要素である。流路を完全に閉塞するためには、バルブ部分の流路断面を半円形として、ダイアフラムの密着性を向上させる必要がある。この構造は一般的に、フォトリソグラフィで作製した矩形断面のポジ型フォトレジストを加熱し、溶融した際に表面張力で形状が円形になることを利用して作製される[88-92]。しかし、基板に塗布できるポジ型フォトレジストの厚さは、約数十 μm が限界である[93]。そのため、培養された細胞組織などの直径数百 μm に達する物体の輸送や大きな流量を発生させるポンプの実装には適していない[94, 95]。

そこで、本章では、厚膜塗布可能なネガ型フォトレジストで作製した大型の流路構造を鋳型として熱可塑性樹脂を成形し、加熱することで、ポジ型フォトレジストでは困難な、大型の半円形流路構造の作製手法を提案する。マイクロ流体デバイスの構造材料として一般的な PDMS で作られた流路に溶融した熱可塑性樹脂を、Vacuum-assisted Micromolding (VAM) [96-100]を応用し、真空力で注入することで流路構造を転写する。加熱時間を変えて加熱を行い、流路断面形状の変化と形状の再現性を調べた。その後、マイクロ流体システムで要求される複雑な曲線構造の作製能力について直角の流路を用いて評価した。さらに、大きさを変えたバルブを作製し、シール能力などを評価した。

3.2 理論、実験方法

3.2.1 EVA ホットメルト接着剤のパターニングとリフロープロセス

図 3.1 に VAM の概要を示す。SU-8 マスターモールドを用いて鋳造した PDMS モールドを、200°Cに設定したホットプレート上のシリコンウェハに貼り付けた。一方のリザーバーに固形 EVA 接着剤(K-13P-12, T-098, Daiso Industries Co., LTD., Japan)を入れ、数十秒放置して溶融させた。毛細管力による流路への自然流入は起こらないため、反対側のポートを減圧して PDMS モールドの流路にグルーを注入した。減圧はダイヤフラムポンプ(DAP-6D, ULVAC KIKO Inc.)を用いて、圧力は約-90 kPa 以上とした。グルーの注入が完了した時点でダイヤフラムポンプを停止し、ウェハと PDMS モールドを冷却した。PDMS モールドはホットプレートの熱で膨張した状態でウェハに接着しているため、冷却時の収縮でウェハから自然剥離し、グルーが硬化前に損傷する可能性がある。そこで、冷却直前に PDMS モールドの上面に PET シート(VF-1410N, KOKUYO Co., Ltd.)を貼り付け、収縮による剥離を防止した。ウェハと PDMS モールドを室温まで冷却した後、PDMS モールドを剥離した。EVA グルーは、シリコンウェハの表面には強い接着力を示すが、PDMS にはほとんど接着しない。この性質を利用することで、成形したグルーを PDMS モールドから容易に離型することができる。VAM で作製した EVA グルーモールドは、SU-8 の流路構造が転写され、矩形の断面形状である。このグルーモールドをホットプレート上でリフローさせると、断面が丸みを帯びた流路が形成できる。

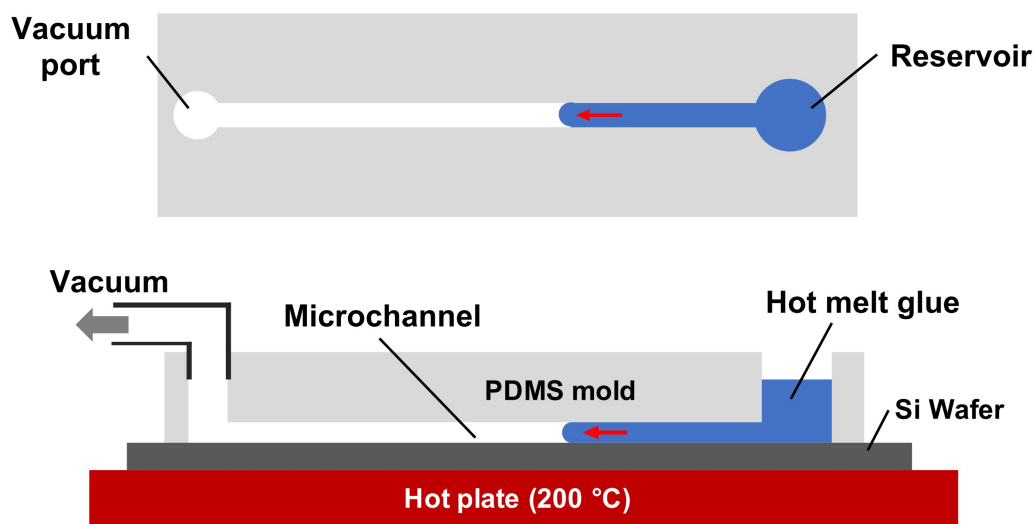


図 3.1 Vacuum-assisted micromolding の模式図

3.2.2 モールドとデバイスの作製

図 3.2 に提案したプロセスを示す。PDMS モールドは、SU-8 マスターモールドから作製した (図 3.2 (a-c))。SU-8 マスターモールドは、ネガ型フォトレジスト(SU-8 3025, Nippon Kayaku Co., Ltd., Japan)をフォトリソグラフィーでシリコンウェハ上にパターンニングすることにより作製した。シリコンウェハ上に SU-8 をスピコートし、95°Cで 60 分間ソフトベークした。このウェハに、流路パターンが印刷されたフォトマスクを介して 250-500 mJ/cm²の紫外線を照射した。その後、95°Cで 5 分間の加熱を行った後、SU-8 developer (Nippon Kayaku Co., Ltd., Japan)を用いて現像を行った。PDMS モールドおよびデバイスは、ソフトリソグラフィーを用いて作製した[86]。SU-8 モールドに厚さ 2 mm のシリコンゴムシート (6-612-04, AS ONE Corporation, Japan) をスペーサーとして貼り付け、デバイスの厚さを調整した。PDMS (Sylgard184, The Dow Corning, USA) を 10 : 1 の割合で混合し、モールド上に塗布した。真空チャンバー内で 30 分間脱泡した後、75°Cのオーブンで 90 分間加熱して硬化させた。硬化後、PDMS を剥離・切断し、直径 4mm のパンチでグルー投入のためのリザーバ穴を打ち抜いた。なお、本章で使用した PDMS モールドはハードベッキングは行われなかった。これは、加熱による変形でウェハと流路端の接着力が低下した場合、熔融したグルーが PDMS モールドの流路から漏れる可能性があるためである。PDMS モールドから VAM を用いてグルーモールドを作製した (図 3.2 (d))。85 °Cに加熱したホットプレート上でモールドをリフローすることで、断面が丸い流路を作製した (図 3.2 (e-f))。断面の形状変化の速度はリフロー温度に比例して大きくなるため、形状制御のための加熱時間の誤差の影響は温度が上がるほど大きくなる。そこで、断面形状の再現性を確保するために、リフロー温度はグルーの熔融温度付近に設定した。その後、グルーのモールドから PDMS をキャストし、デバイス作製と評価を行った (図 3.2 (g-h))。

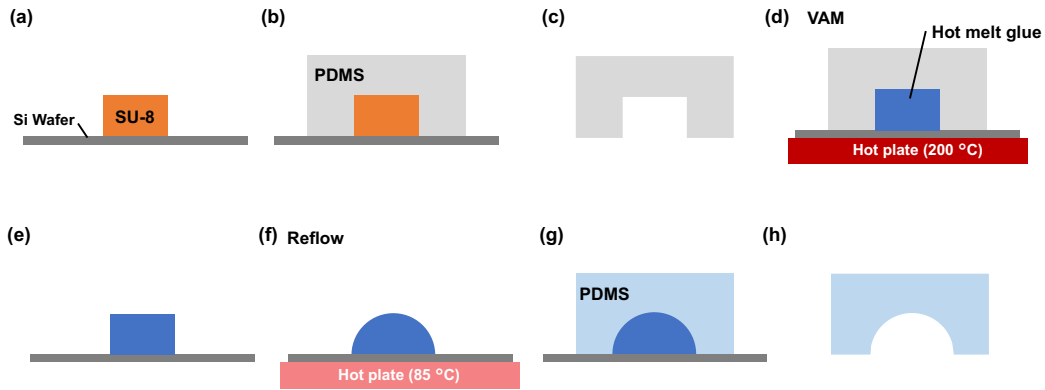


図 3.2 VAM を用いたモールド作製とリフロープロセスの手順。(a)フォトリソグラフィでネガ型フォトレジストのマスターモールドを作製する。(b-c)マスターから VAM のための PDMS モールドを作製する。(d)VAM による EVA ホットメルト接着剤のパターニング。(f)グルーモールドをホットプレートで加熱し、流路断面形状を。(g,h)リフロー後のグルーモールドで PDMS を鋳込む。

3.2.3 流路断面とパターニング能力の評価

流路の断面形状を評価するために、幅 100、300、800 μm 、長さ 40 mm の直線流路を使用した。SU-8 マスターモールドは前節で述べた手順で作製した。上記の流路に対応する 3 つのモールドのレジスト厚さはそれぞれ 48.8、155.1、494.1 μm であった。PDMS モールドは SU-8 マスターモールドから作製し、VAM モールドと PDMS モールドを用いてシリコンウエハ上に EVA グルーをパターニングした。その後、グルーモールドをホットプレート上でリフローし、PDMS ソフトリソグラフィを用いてレプリカを作製した。レプリカの断面を顕微鏡で観察し、PDMS レプリカの断面像の幅 w と高さ h から、流路の接触角を算出した。接触角 θ は以下の式で算出した。

$$\theta = 2 \tan^{-1} \left(\frac{h}{w/2} \right) \quad (3.1)$$

図 3.3(a)は、パターン形成能力を示すために使用した流路構造である。同一基板上に曲率半径の異なる流路パターンを作製し、リフローを行った。曲率半径は流路幅 w の設計値の定数倍である $r_n = nw$ とした。SU-8 マスターモールドのレジスト厚さは 49.6 μm と 421.4 μm であった。リフローしたグルーモールドで作製した PDMS レプリカの流路幅を測定し、カーブパターンを評価した。リフローの際、グルーの溶融・変形により湾曲部の内側のエッジが湾曲部の内側に移動していることがわかった。このとき、流路の断面積は保存されていると

仮定すると、流路の高さは減少していることになる（図 3.3 (b)）。曲線流路の形状変化は、直線と曲線流路の幅の差を測定することで評価した。

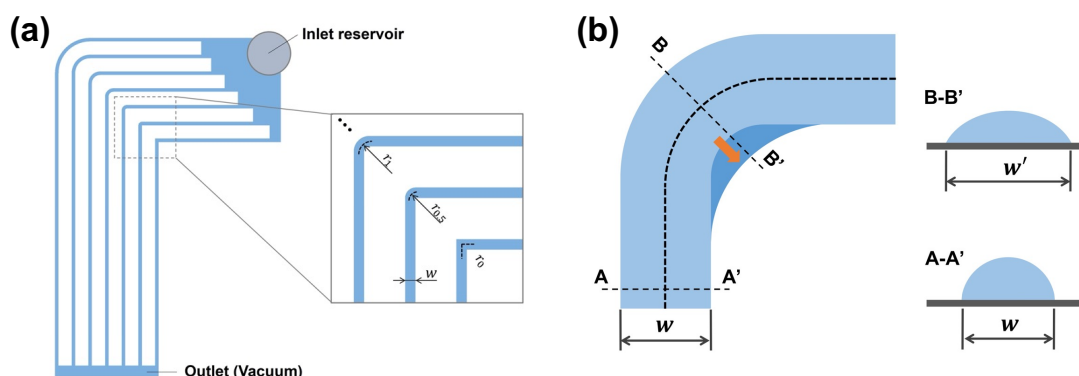


図 3.3 (a)モールドの流路構造。(b)曲線の評価方法。

3.2.4 バルブシール性能とポンプ流量の測定

図 3.4 は、マイクロペリスタリックポンプの構造を示している。このポンプは、channel layer、ダイヤフラム、pneumatic layer の 3 層で構成されている。マイクロ流路の設計幅は 100、300、800 μm 、pneumatic チャンバーの幅 w_p はそれぞれ 250、800、1500 μm であった。これらの直線流路は、VAM を介してシリコンウェハー上に 100、300、800 μm の幅でパターンニングし、85°C のホットプレート上でそれぞれ 30、60、300 秒間、リフローを行った。モールド製作後、channel layer のモールドには厚さ 3mm、pneumatic layer のモールドには厚さ 2mm のシリコンゴムシートをスペーサーとして貼り付け、デバイスの厚さを調整した。その後、PDMS モールドと同じ工程で、PDMS をキャストした。モールドから剥離した PDMS の外形を切断し、直径 8mm のパンチでそのチャンバーを切断した。基剤と硬化剤を 10 : 0.7 の割合で混合した PDMS をシリコンウェハー上にスピコートしてダイヤフラムを作製した。これを 75°C で 90 分間加熱することにより硬化させた。ダイヤフラムの厚さは 26.2-28.5 μm であった。ダイヤフラムと pneumatic layer は空気プラズマを照射し、ホットプレート上で 75°C、15 分間加熱して接着した。この層と channel layer を同じ方法で接着した。マイクロポンプの加圧システムは、ダイヤフラム駆動用の電磁弁 (M3QB110-M5-C2H-8-4, CKD, Japan) と Arduino マイクロコントローラ (Arduino Uno, Arduino SRL, Italy) から構成された。Pneumatic チャンバーにはダイヤフラムのたわみを防ぐため Milli-Q 水を満たした。バルブの蛍光観察は顕微鏡 (IX71, Olympus Corporation, Japan) に取り付けられた sCMOS カメラ (ORCA-flash4.0, C11440-22CU, Hamamatsu Photonics K.K.) を用いた。マイクロ流路には、Milli-Q 水で 1 mM に希釈したフルオレセイン (F6377-100G, Sigma-

Aldrich Co., USA) を充填した。幅 $100\ \mu\text{m}$ の流路を観察する場合は、10 倍対物レンズ (UPLANFLN10x, Olympus Corporation, Japan)、ミラーユニット (U-MNIBA2, Olympus Corporation, Japan)、フィルター (25ND25, Olympus Corporation, Japan) を使用した。その他の流路には、4 倍対物レンズ (UPLANFL4x, Olympus Corporation, Japan)、ミラーユニット (U-MWIB2, Olympus Corporation, Japan) を使用した。流速測定時には、Tween 20 (170-6531, Bio-Laboratories, Inc., USA) と Victoria Blue B (228-00222, Wako Pure Chemical Industries Ltd., Japan) をそれぞれ 0.05% (v/v) および 0.1 mM に Milli-Q water で希釈したものを流路に注入した。ポリエチレン粒子 (REDPMS-1.085 27-32 μm , Cospheric Microspheres, USA) を $w=100\ \mu\text{m}$ と $300\ \mu\text{m}$ の装置で流体トレーサーとして使用した。 $w=800\ \mu\text{m}$ の装置では、90-106 μm の粒子 (REDPMS-1.065 90-106 μm , Cospheric Microspheres, USA) を配置した。流量は、ポンプ作動中の粒子を撮影し、流路の断面積と移動した粒子数から算出した。バルブの開閉状態を open(0)、close(1)と表記するとき、使用した動作シーケンスは(001)(100)(111)である。動作周波数は、バルブの 1 状態遷移周期を示す。

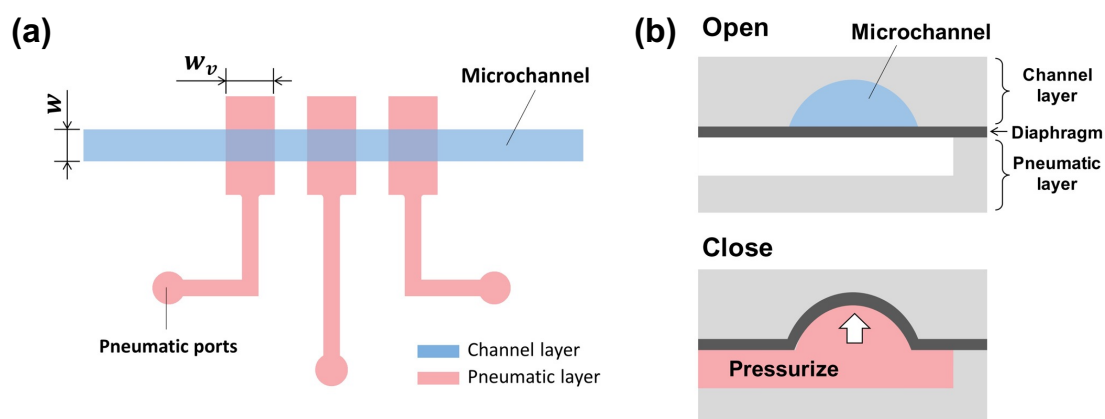


図 3.4 (a) マイクロペリスタポンプの構造 (b) バルブの断面図

3.3 実験結果および考察

3.3.1 形状再現性評価

流路の形状を評価するために、流路と基板間の接触角を測定した。議論の便宜上、断面積の小さいものから順に、Type A、B、C と表す。減圧開始から一定距離の流路にグルーが充填されるまでの時間は、断面形状、グルー粘度、減圧圧力によって決まる[96]。最も大きい流路 (Type C) は $40\ \text{mm}$ の流路を充填するのに、数秒程度であるが、最も小さい流路 (Type A) は $30\ \text{mm}$ の充填に約 10 分必要である。流路断面の測定位置は、成形された流路の中心付近の寸法で測定した。すなわち、 $300\ \mu\text{m}$ と $800\ \mu\text{m}$ の流路は注入口から $20\ \text{mm}$ 、 $100\ \mu\text{m}$ の流路は注入口から $15\ \text{mm}$ である。グルーを充填する流路長さが長く、直径が細くなるほ

ど流路抵抗が大きくなり、グルーの流速が低下する。経験上、等価直径 $30\ \mu\text{m}$ 程度の流路では、 $30\ \text{mm}$ の長さを充填するのに約 2 時間必要となる。また、流路幅が $30\ \mu\text{m}$ 未満の場合、リフロー時の濡れ挙動の安定性に問題があるケースもあった。パターンニング後、グルーモールドを 85°C のホットプレートで加熱した。

図 3.5 に加熱前と加熱後の Type C のモールドを示す。このように、リフローにより断面が丸みを帯び、表面が滑らかになっていることが確認できた。図 3.6 (a,c-h) は、グルーモールドを用いて作製した PDMS レプリカの断面図である。この実験では、PDMS の硬化温度をグルーの融点付近に設定したが、 75°C の加熱でも流路の形状は変化しなかった。これは、直方体のグルーモールドを用いて PDMS を複数回キャストしても、流路のエッジや壁の粗さが維持されることから確認された。ただし、加熱を行う環境によっては、形状が変化する可能性がある。その場合、 $5\sim 10^\circ\text{C}$ の低温で加熱し、PDMS の硬化時間を長くするなどの対策が考えられる。図 3.6 (a,c) はリフロー前の SU-8 モールドとグルーモールドの断面図である。断面形状を定量的に評価するために、ウェハに対する流路の接触角を用いた (図 3.6(b))。図 3.6 (i) は 3 種類のモールドについて、加熱時間に対する接触角の推移を示したものである。また、表 3.1 に各加熱時間における流路寸法と接触角の推移を示す。ここで、表中の $0\ \text{s}$ は、直方体のモールドの寸法である。初期の矩形断面からリフロー時間の増加とともに接触角が減少し、最終的に約 50° に近づくことが確認された。流路の断面が大きくなると、相転移に必要な熱エネルギーが増加し、流路構造の上面がホットプレートの加熱面から遠くなる。そのため、流路の形状変化は緩やかに進行する。この実験で得られた流路の最大高さは $503\ \mu\text{m}$ で、この流路の幅は $996\ \mu\text{m}$ であった。3 種類の流路において、接触角の変化を示す変動係数は最大で約 4.2% であり、再現性の高いプロセスであることが確認された。このことから、加熱時間を変化させることで、モールドの断面形状を制御することができることが示された。また、5 回以上の PDMS 鋳造を行った場合でも、形状変化やグルーモールドの破損は確認されなかった。さらに、VAM を 5 回以上繰り返した場合でも PDMS モールドの破断は観察されなかった。

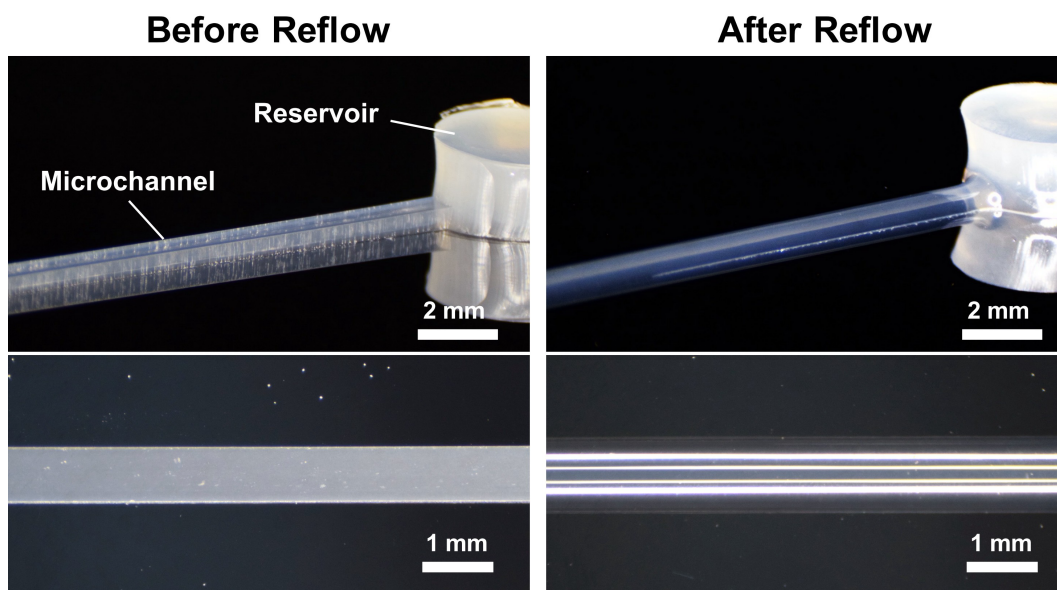


図 3.5 リフロー前後のグルーモールド

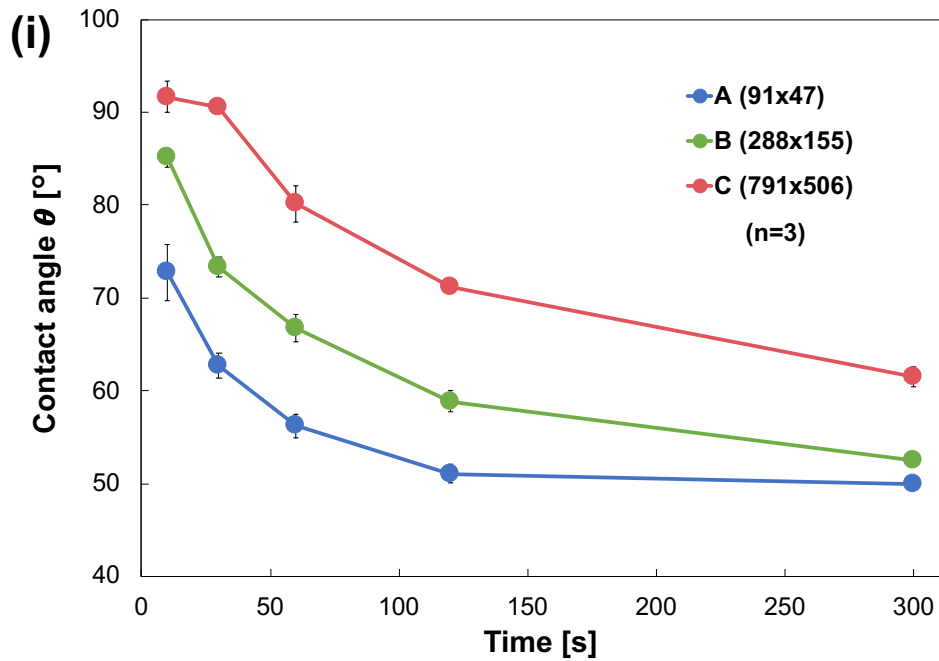
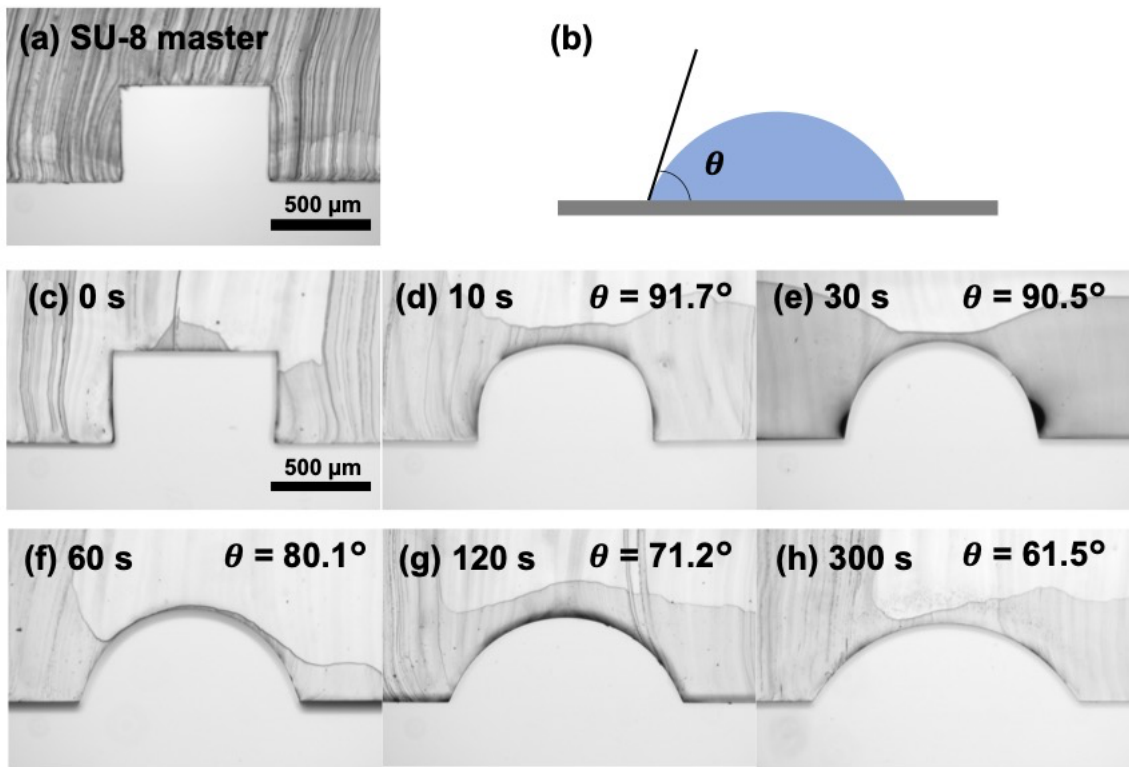


図 3.6 モールドの断面形状 (a) SU-8 マスターモールド。(b)流路接触角の定義 (c) リフロー前のグルーモールド (d-h) 10~300 秒のリフローで作製したモールド。(右上に示す接触角は 3 回の実験の平均値である)。(i) 流路の接触角とリフロー時間。凡例の数字はリフロー前のグルーモールドの矩形流路の寸法 (幅×高さ) [μm]を表す。エラーバーは標準偏差を示す。

表 3.1 各加熱時間における流路寸法と接触角。誤差の範囲は標準偏差を示す。

Time [s]	A (96x49)			B (288x155)			C (792x494)		
	Width [μm]	Height [μm]	θ [$^\circ$]	Width [μm]	Height [μm]	θ [$^\circ$]	Width [μm]	Height [μm]	θ [$^\circ$]
0	97.8 \pm 1.0	43.3 \pm 0.1	-	299.8 \pm 1.0	139.5 \pm 0.3	-	829.1 \pm 5.7	467.4 \pm 2.5	-
10	126.4 \pm 4.2	46.5 \pm 1.0	72.7 \pm 3.0	342.3 \pm 10.8	157.2 \pm 2.5	85.2 \pm 1.0	881.2 \pm 14.8	492.9 \pm 7.1	96.4 \pm 1.6
30	139.2 \pm 3.4	42.4 \pm 0.1	62.7 \pm 1.3	388.2 \pm 4.1	144.6 \pm 1.9	73.4 \pm 1.1	984.2 \pm 13.3	496.7 \pm 5.2	90.5 \pm 0.6
60	149.4 \pm 2.0	39.9 \pm 0.6	56.2 \pm 1.3	418.7 \pm 5.7	137.9 \pm 2.2	66.8 \pm 1.5	1109.0 \pm 26.3	466.4 \pm 5.7	80.1 \pm 2.0
120	157.9 \pm 2.9	37.7 \pm 0.3	51.1 \pm 1.0	452.5 \pm 9.6	127.6 \pm 0.4	58.9 \pm 1.2	1195.2 \pm 20.3	427.9 \pm 7.1	71.2 \pm 0.6
300	160.0 \pm 0.8	37.2 \pm 0.6	49.9 \pm 0.5	488.4 \pm 1.7	120.3 \pm 0.7	52.5 \pm 0.2	1335.1 \pm 14.6	397.5 \pm 5.2	61.5 \pm 1.1

3.3.2 パターニング能力評価

パターニング能力を実証するために、直角に曲がる流路構造を作製した。リフロー後の湾曲部と直線部の幅の差を測定することで、湾曲部に断面形状が保持されているかを確認することができる。図 3.7(a,b)は幅 800 μm のモールドのリフロー前とリフロー後の様子である。隣接する流路へのグルーの漏れは観察されなかった。図 3.8 は曲率半径に対する湾曲部と直線部の流路幅の誤差を示したものである。曲率半径が $w/2$ の場合、曲線部の幅は直線部の幅に比べて 20-30%増加した。このことは、流路の断面形状が大きく変化したことを示している。フィレット半径 $r = 0$ の流路（図 3.7 (a,b)の左端の流路）では、リフロー後の直線部の幅の $\sqrt{2}$ 倍が理想の幅と仮定して、 $w = 100$ と 800 でそれぞれ 0.65%、4.25% の誤差が生じた。フィレットがない流路の場合の誤差が比較的小さいのは、コーナーの外形線が内側に進展し、誤差が打ち消されたためと考えられる。図 3.8 より、曲率半径が大きくなるにつれて、曲線部での誤差が小さくなっていることがわかる。実際に、半径が初期流路幅の 3 倍以上になると 5%以下となった。この結果から、十分な曲率を持つ流路では、湾曲部の断面が大きく変形することはないことがわかった。したがって、適切なフィレット半径を設定することで、本手法により湾曲部の形状が大きく変化しない複雑な流路構造を作製することが可能である。

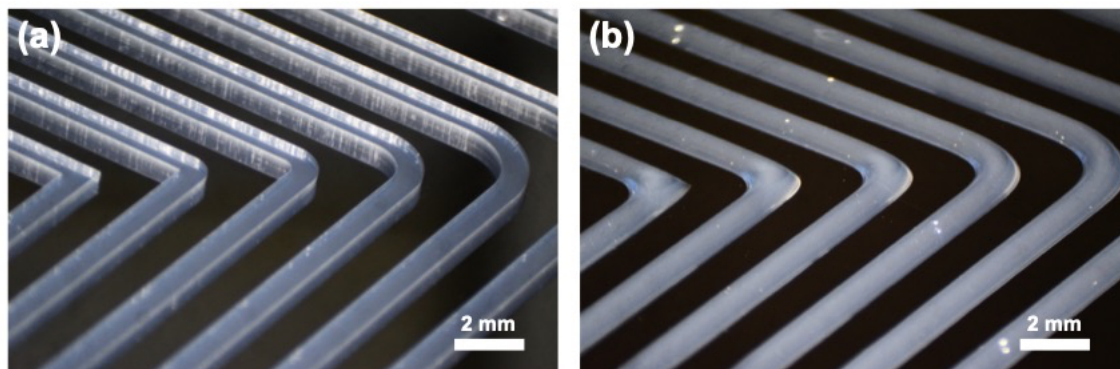


図 3.7 曲線評価用デバイスのグルーモールド (a) リフロー前 (b) リフロー後

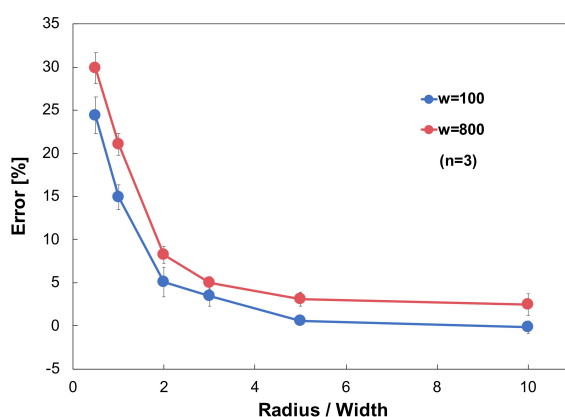


図 3.8 直線部の流路幅に対する曲線部の幅の増加割合。エラーバーは標準偏差を示す

3.3.3 バルブおよびポンプの性能評価

本手法のマイクロ流体デバイスへの応用として、3つのマイクロバルブからなるマイクロペリスタルティックポンプを作製し、その性能を評価した。マイクロバルブとしての流路の密閉性能を示すため、各デバイスの流路内に蛍光色素溶液を入れ、バルブ開閉時の蛍光強度を観察した。図 3.9 (a) は、バルブの開閉状態での蛍光観察と明視野観察の画像である。空気圧チャンバーを加圧し、ダイヤフラムが上昇すると、バルブ内の蛍光色素溶液は排斥され、蛍光強度はバックグラウンドに近づく。これは、流路が完全に閉鎖されたことを示している。図 3.9 (b) は、チャンバーに加えた圧力に対するバルブの開閉度を示している。測定領域内の蛍光強度の平均値が開閉度として示されている。断面積の異なる全てのタイプのバルブが 60 kPa 以下で閉塞することを確認した。

図 3.10 にポンプの流量の測定値を示す。流路内の微粒子をトレーサーとして流量を測定した。どのデバイスも 20 Hz 程度までは動作周波数に比例して流量が増加し、その後、増加率は徐々に減少した。最大流速は 50 Hz で 237 $\mu\text{L}/\text{min}$ であり、これは流路寸法の最も大き

い Type C の装置で観察されたものである。この節で説明したバルブとポンプは、高流量を必要とするシステムや、細胞組織のような大きな粒子の輸送に適用可能である。

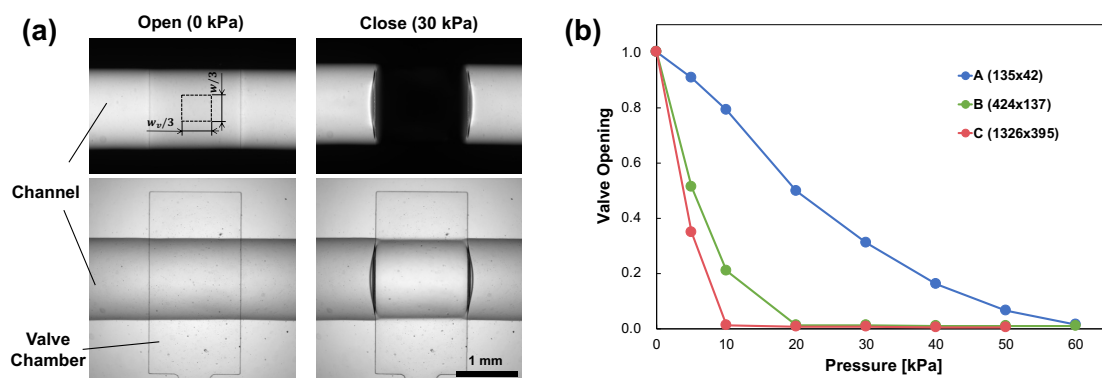


図 3.9 (a) バルブ開閉時の様子。上段は蛍光観察、下段は明視野観察の画像。左上の画像の破線の領域は、蛍光強度の測定領域（バルブ中心から w_v と w の 1/3 の領域）。(b) バルブの相対的な開閉度。0 kPa の圧力を 1.0 としたときの測定領域における平均強度を示す。凡例中の数字は半円形流路の寸法（幅×高さ） [μm] を表す。

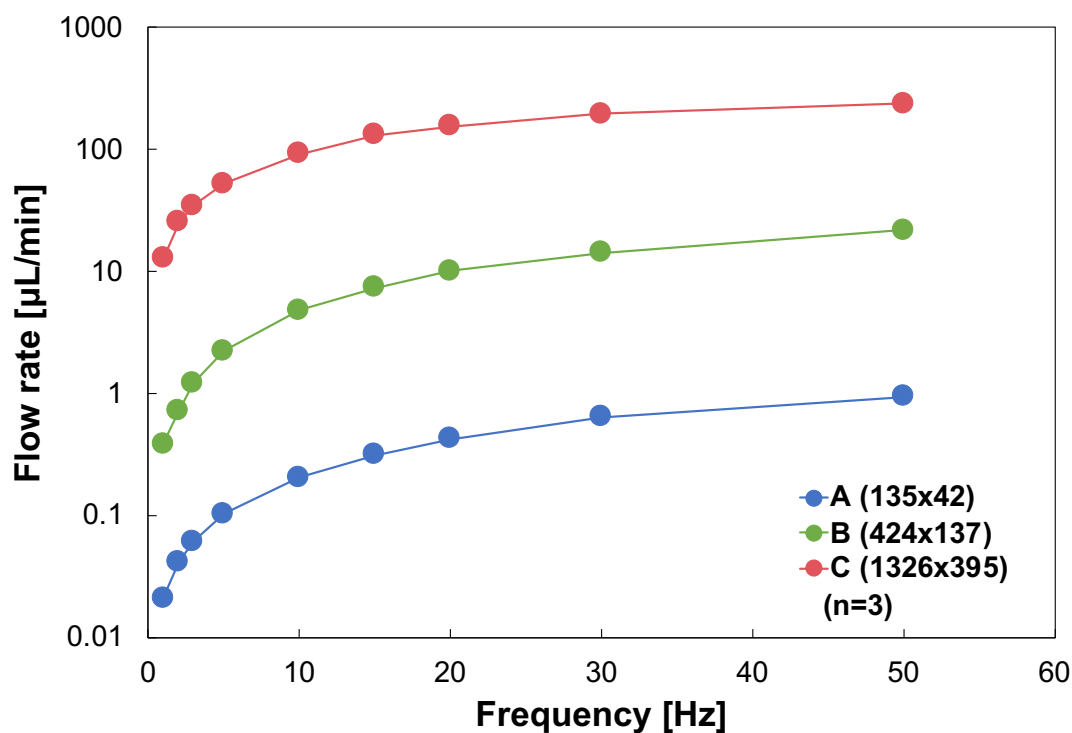


図 3.10 マイクロペリスタルティックポンプの動作周波数に対する流量の測定結果

3.4 結言

本章では、EVA ホットメルト接着剤と VAM を用いたリフロープロセスを報告した。この方法は、SU-8 マスターモールドから作製した PDMS モールドを用いて、EVA グルーを基板上にパターンニングすることにより、数百 μm 規模の構造を持つ半円形流路断面を持つモールドを作製することができる。まず、VAM を用いてシリコンウェハ上に直線流路をパターンニングし、リフロープロセスを実行した。作製したモールドを用いて PDMS をキャストし、断面を観察したところ、丸みを帯びた断面が形成された。各加熱時間における形状変化の CV は 5% 以下であり、加熱時間により断面形状を制御できることが確認された。また、本実験で得られた半円形流路の最大高さは、幅 996 μm のとき 503 μm であった。続いて、直角に曲がる流路構造を作製し、複雑なパターンの形成能力を評価した。グルーの形状変化によるパターンの著しい変形は観察されなかった。本手法の応用例として、3つのバルブからなるマイクロペリスタルティックポンプを作製した。バルブの蛍光観察やポンプ流量の測定から、本手法がマイクロ流体デバイスの作製に適用可能であることが示された。本手法は、比較的大規模な流路構造を作製することが可能であるだけでなく、切削加工よりも滑らかな型面を得ることができる。また、PDMS や EVA のモールドは繰り返し使用することができる。加えて、ホットプレートと小型真空ポンプで行うことが可能で、PDMS モールドへの表面処理も不要である。したがって、数百マイクロメートルスケールの構造体を低コストで簡便に作製するプロセスとして利用することができる。

第4章 2次元微粒子操作技術の開発

4.1 緒言

本章では、前章で確立した粒子操作技術を2次元平面上への操作へ拡張する。粒子を操作する領域を取り囲むように流路とバルブを配置し、それらを制御することで粒子の位置を2次的に操作するデバイスを設計した。

ここで、強化学習は試行錯誤によって最適な行動を獲得するため、実環境のみでの学習は膨大なコストを要する。そのため、本章においても、学習の効率化のためシミュレータを利用する。粒子操作タスクは、バルブ動作に対する粒子の移動を予測できれば、シミュレートできる。しかし、数値流体力学に代表される流体の計算は、大きな計算コストを必要とするため、高速化が難しい。そこで、本章では、液中微粒子の挙動の予測技術として、システムの学習前にCFDソフトウェアで流路内部の速度データ取得し、これを粒子の速度として2次元の物理エンジンに入力することで粒子の移動を高速に予測する手法を提案した。この手法による予測結果と実測した粒子の運動を比較し、本手法で液中粒子の移動予測が可能であることを示した。

シミュレータの能力を確認した後、このシミュレータを用いて2次元平面における粒子操作の学習を行い、そのモデルを実環境にデプロイすることで新しい2次元粒子操作技術を実証した。

4.2 理論、実験方法

4.2.1 デバイスの構造と強化学習の適用

本システムへの強化学習の適用のため、タスクの設定と特徴量の定義を行った。マイクロバルブは Open/Close の 2 状態のみを取ることが可能で、Open/Close の状態の組合せでバルブのフェーズを表す。現在の開閉状態から、バルブの開閉が完全に遷移し、粒子の位置を観測するまでを単位タイムステップとした。ここで、粒子操作タスクにおける状態と行動は以下のように定義した。

State s : Position of beads p_b , Position of targets p_t , Current phase of valves v

Action a : Next phase of valves

報酬は、タイムステップ t でバルブを操作して粒子がターゲット位置に到達した場合に $r_t = 10$ が与えられ、それ以外は $r_t = 0$ とした。また、粒子の操作はデバイスの操作領域内のみで行われ、マイクロ流路部分に粒子が出た場合は、その時点でタスクを終了し次のエピソードに移行する。

強化学習のアルゴリズムは Double Deep Q Network (DDQN) [101] を用いた。状態 s_t における行動 a_t の行動価値の推定値を、policy ネットワークのパラメータ θ_t を用いて $Q(s_t, a_t; \theta_t)$ と表すとき、タイムステップ t のネットワークの損失関数は、

$$L_t(\theta_t) = \mathbb{E} \left[(Y_t - Q(s_t, a_t; \theta_t))^2 \right] \quad (4.1)$$

である。ここで、教師信号 Y_t は、一定タイムステップの間固定されるパラメータ θ_t^- を用いて、

$$Y_t = r_t + \gamma Q \left(s_{t+1}, \underset{a}{\operatorname{argmax}} Q(s_{t+1}, a; \theta_t); \theta_t^- \right) \quad (4.2)$$

である。タイムステップ t におけるエージェントの行動選択は、以下の戦略に従う。

$$a_t = \underset{a}{\operatorname{argmax}} Q(s_t, a; \theta_t) \quad (4.3)$$

学習は experience replay[64] が利用された。これは、タイムステップ t における経験 $\langle s_t, a_t, r_t, s_{t+1} \rangle$ をバッファに蓄積し、ネットワークの学習時はこのバッファからランダムにバッチサイズ分の経験をサンプリングし、そのデータセットを用いてネットワークを更新する手法である。

policy ネットワークは 512 unit の隠れ層を 1 層持つ Multi-layer perceptron (MLP) で構成される。隠れ層の活性化関数は Rectified Linear Unit (ReLU) function が用いられ、optimizer は Adam [102] が用いられた。学習開始後 10000 step はランダムな行動で経験をバッファに蓄積し、この間はネットワークの学習は行われない。ランダムアクションで経験を蓄積した後、 ϵ -greedy 方策を用いて学習が行われ、100 episode までは $\epsilon=1.0$ から 0.1 まで減少させ、その後は 0.1 で固定した。ターゲットネットワークのパラメータ θ_t^- は 10000 step で最新のタイムステップのパラメータ θ_t に更新した。Replay memory のキャパシティは 10^5 step とした。

4.2.2 粒子操作シミュレーションおよび実験環境

デバイスの構造を図 4.1(a)に示す。中心部の操作領域(Operating area)内の粒子の挙動を予測する。この領域につながる流路に、マイクロバルブと呼ばれる機構を配置する。これは、空気圧を利用して薄膜を動かすことで流れを生じる機構である。各バルブの Open/Close の状態を切り替えることで操作領域内の粒子の位置を制御する。図 4.1(b)は学習システムの概要である。デバイスの操作領域をカメラで撮影し、コンピュータで画像処理を行い、粒子位置を取得する。それと同時にバルブの開閉状態の制御も行う。

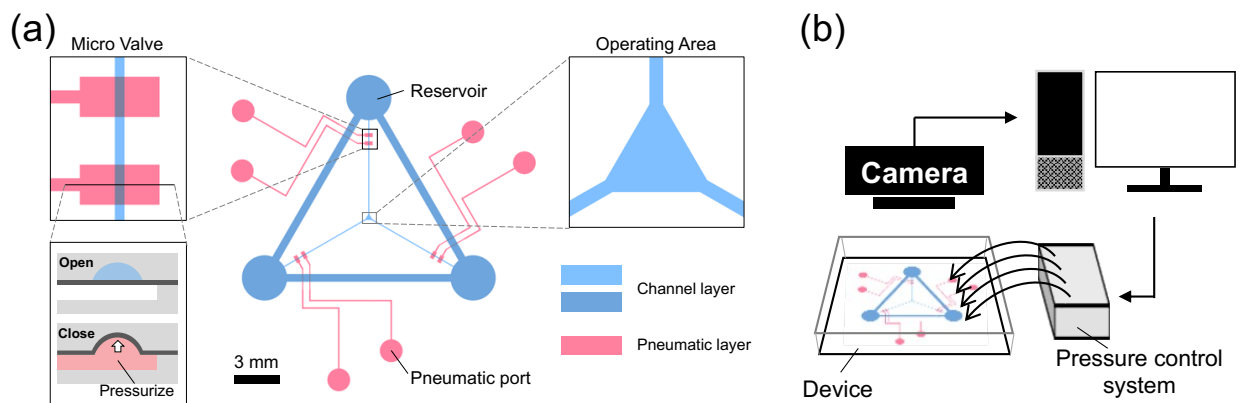


図 4.1 (a)デバイスの構造 (b)学習システム

4.2.3 CFD と 2D 物理エンジンを利用した液中粒子挙動の予測

図 4.2 に提案手法の概要を示す。マイクロバルブの動作で生じる流体の移動を操作領域につながる流路で測定し、これを境界条件として領域内の流速を CFD ソフトウェア (COMSOL Multiphysics)を用いて予測する。すべてのバルブ動作に対して流速を計算し、速度場のデータを粒子の速度として参照し、2次元の物理エンジン (Pymunk, <http://www.pymunk.org>)でその後の運動を予測する。マイクロバルブで生じる流れは薄膜の変形量に依存するため時間変化するが、本研究ではバルブの開閉状態が切り替わるまでに移動した体積が単位時間あたりに流れると仮定することで 3次元定常非圧縮性流れとして扱った。また、領域高さ h に対して $z = h/2$ の平面の流速を粒子速度として使用した。物理エンジンでのシミュレーションは、1回のバルブ動作あたり、2.5 ms のタイムステップで 500 ms 間実行され、その間、粒子の速度を 50 ms ごとに、位置情報をもとに更新した。直径 $20 \mu\text{m}$ の粒子を用いて実験を行った。

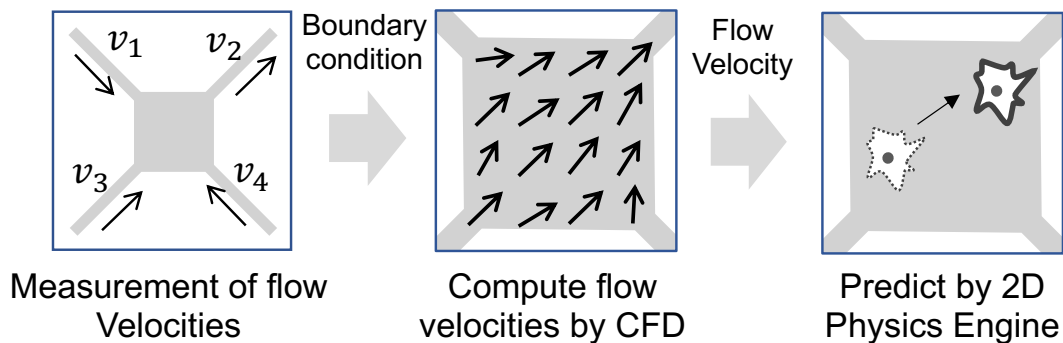


図 4.2 マイクロ流路内の粒子運動シミュレーション手法

4.2.3 デバイスの製造

前章で実証した VAM を用いたマイクロバルブの製造方法を応用してデバイスを作成した。前章の VAM によるリフロープロセスでは、Channel layer を直接減圧したが、このデバイスでは外周部と内部の流路直径が大幅に異なるため、内部にグルーが十分に充填されないことが想定された。そのため、図 4.3 に示すように、デバイスの Channel layer に隣接する形で真空用の流路を設けた。各リザーバにグルーを投入し、Vacuum layer の流路を減圧すると、PDMS のガス透過性によって Vacuum layer に隣接する Channel layer も減圧され、グルーが Channel layer 内部に充填される。この手法を利用して、SU-8 で構築された流路パターンを基板に転写した。

モールドの製造方法は以下の通りである。まず、図 4.3 に示すデバイスの Pneumatic layer 以外の流路構造をフォトリソグラフィを用いて SU-8 で作製した。Channel layer の外周部のモールド厚さは $251.83 \mu\text{m}$ 、操作領域を含む内部の厚さは $26.82 \mu\text{m}$ とした。前章までと同様の方法で、ソフトリソグラフィによってグルー成形用の PDMS モールドを作製した。PDMS モールドの変形を抑えるため、今回はグルー成形時のホットプレート温度を 150°C とし、グルーモールドを作製した。リザーバにグルーを投入し、Vacuum layer を減圧することで、グルーをパターンニングした。作製したグルーモールドのバルブ部分のみを半円形断面とするため、図 4.4(a)のようにハンダゴテを用いてグルーモールドの裏側から加熱することで局所的にグルーを溶融した。図 4.4(b)はリフロー後のグルーモールドである。バルブ部分以外の矩形断面を保持した状態で半円形上の流路が得られた。

このグルーモールドを用いてデバイスを PDMS で作製した。PDMS の鋳込みの手順はこれまでと同様である。また、Pneumatic layer は SU-8 で作製したモールドを使用して作製した。ダイアフラムは基剤と硬化剤を 10:0.7 の比率で混合した PDMS をシリコンウェハにスピコートすることで作製した。ダイアフラムの厚さは $19.19 \mu\text{m}$ であった。それぞれのレイヤをプラズマ接合し、デバイスを作製した。

デバイスの流路内を、Milli-Q water に Tween20 (170-6531, Bio-Laboratories, Inc., USA) を 0.05%(v/v)で希釈したもので満たし、デバイスの流路に 20 μm のポリスチレン粒子(#18329, Polysciences Inc., USA)を一つ配置し、操作を行った。

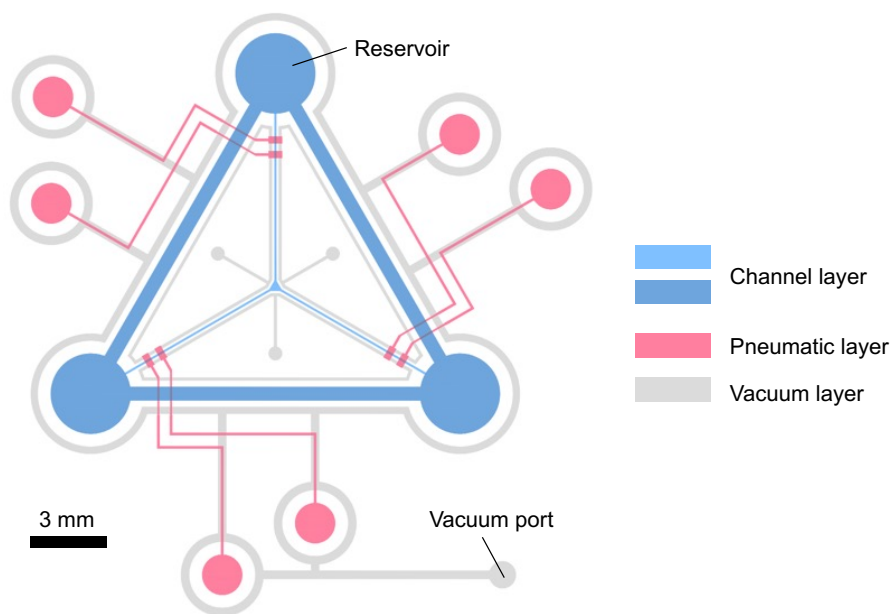


図 4.3 VAM のためのモールドの流路構造

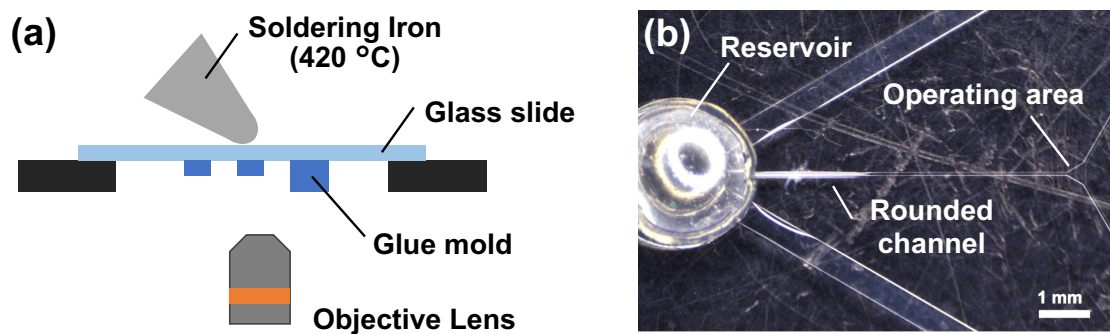


図 4.4 (a)局所的なリフロープロセスの概要 (b)リフロー後のグルーモールド

4.3 実験結果および考察

4.3.1 予測能力の評価

CFD による速度場の予測のため、境界条件として、マイクロ流路部分の流速を測定した。図 4.5 は、任意のバルブ状態から任意のバルブ状態への遷移において、各流路でサンプリングされた流量の合計値である。デバイスの流路内で質量保存則が成り立つので、流路から操作領域への流入と流出は合計して 0 になると考えられる。図 4.5 は色が濃いほど流入出の合計値の誤差が大きいことを示している。ここで、いくつかの遷移で粒子の認識ミスによる値が確認されたため、デバイスの対称性から、相似な遷移の平均値を測定値としてプロットした。全体の誤差は平均 1.5 px 程度となり、デバイス内の水頭圧差によると思われる対流などは確認できなかった。外周の流路によってデバイス内の圧力を一定に保つことができると考えられる。このデータを利用し、COMSOL による速度場の予測を行い、シミュレータを構築した。

続いて、バルブをランダムに動かした場合の実測値と予測値の比較が行われ、誤差を評価した。ここで、事前にバルブをランダムに遷移した場合の領域内の粒子移動量をサンプリングした。この測定では、粒子の初期位置 p_0 と行動 α 実行後の粒子位置 p_1 を記録した。シミュレーションでは、初期位置を p_0 とし、行動 α 実行後の粒子位置 p'_1 を予測した速度場と物理エンジンを利用して算出した。 p_1 と p'_1 から予測精度を評価した(図 4.6)。誤差の評価には、二乗平均平方誤差(RMSE)を用いた。

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_{1n} - p'_{1n})^2} \quad (5.4)$$

これは、誤差が二乗されるため誤差の絶対値が大きなもの重視する。この誤差を最小化するように物理エンジンのパラメータを調整した。

図 4.7 は粒子の移動の実測値と予測値の比較である。バルブをランダムに動かした場合の平均粒子移動量は 24.65 μm であり、この実測値に対する二乗平均平方根誤差は 4.35 μm であった。また、本手法は粒子の挙動予測時には流体の計算を必要としないため、動作速度は実機に比べて 100 倍以上高速であることが確認された。

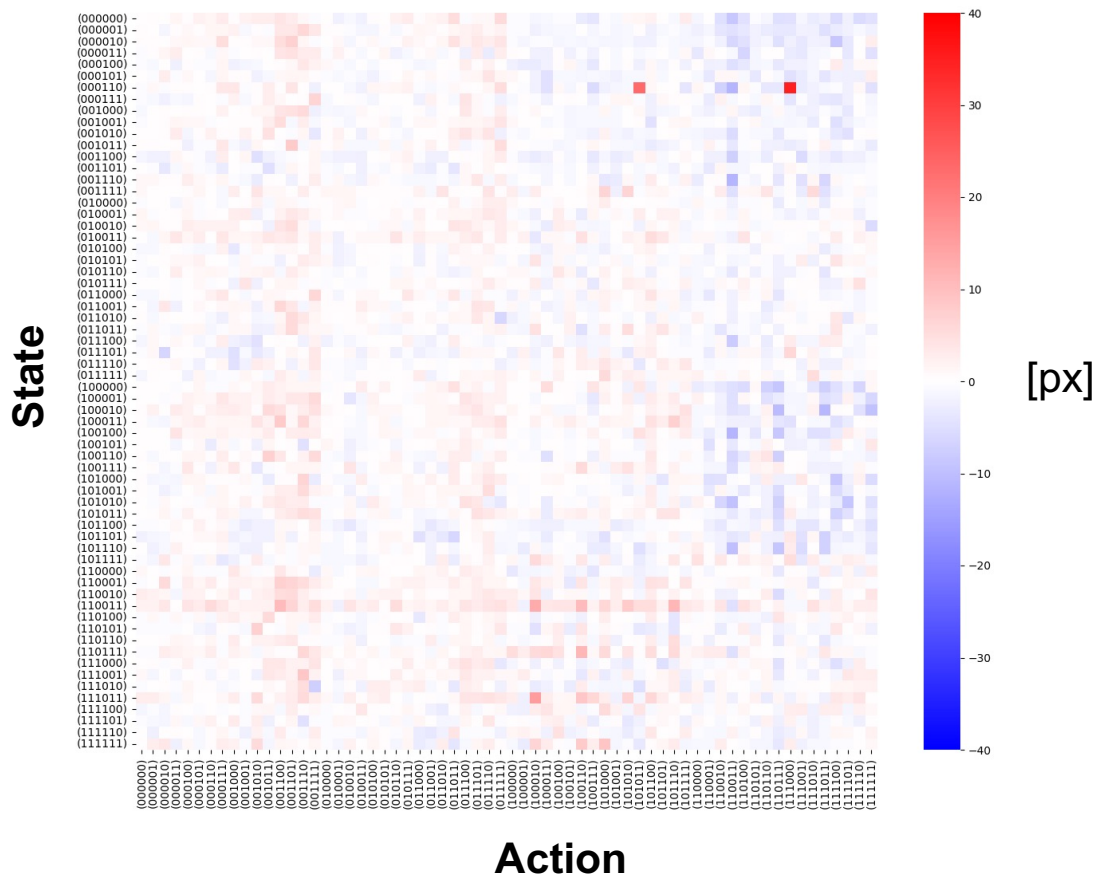


図 4.5 任意の状態行動の組み合わせにおける、3つの流路の粒子移動量の合計値（1px あたり $2.285 \mu\text{m}$ ）

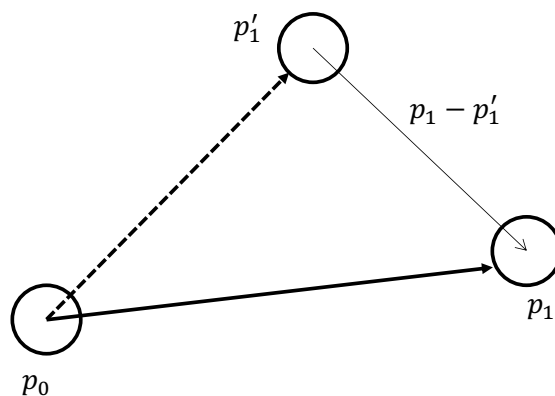


図 4.6 粒子移動量の比較方法

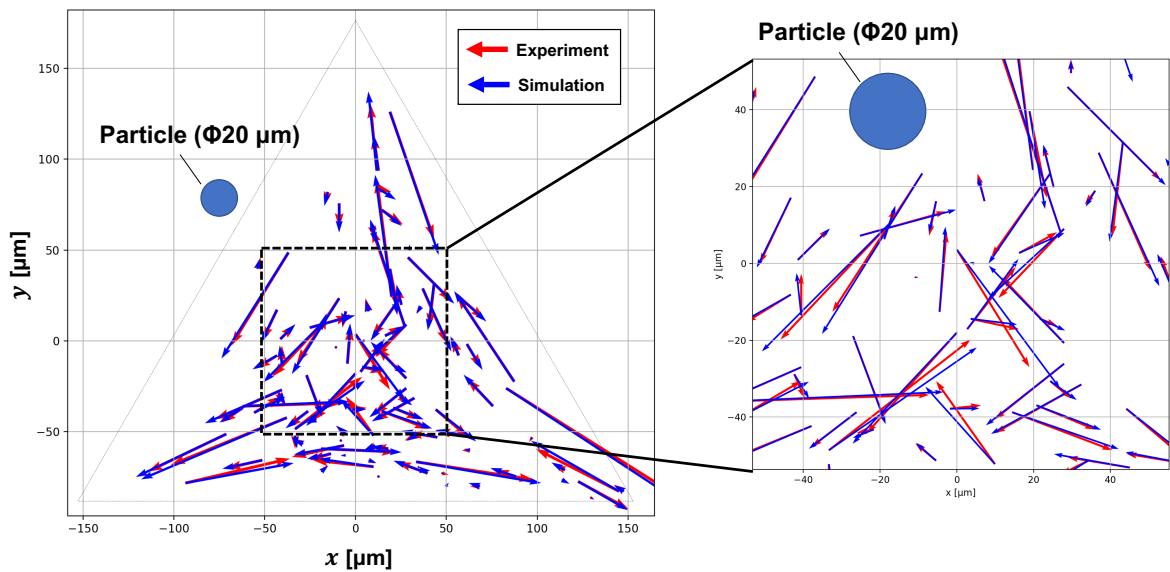


図 4.7 CFD ベースのシミュレータによる粒子移動予測の結果

4.3.2 粒子操作の実証

ターゲットの大きさを 2 種類とし、ターゲット位置をランダムにしてシミュレーションを行った。図 4.8、表 4.1 より、どちらの設定でも、報酬の上昇とステップ数の減少が確認された。ターゲットエリアが大きい方が学習の収束が早く、ステップ数も少ない。シミュレーションでターゲットを小さく設定して学習を行うことで、同じモデルで大きなターゲットにも対応できると考えられる。以上より、学習が正常に行われたと判断し、この 2 つのモデルを実環境でテストした。

表 4.2 より、それぞれのターゲットサイズで、タスク成功率は 98% 以上、ステップ数はシミュレーションと同じ程度であった。ただし、シミュレーションは粒子初期位置がランダムなので、実際の成績は実験よりも高いと考えられる。ターゲット位置をランダムとした場合でも、これまでのワークフローが機能することが示された。また、ターゲットサイズを小さくしてもタスクを実行できることが確認された。

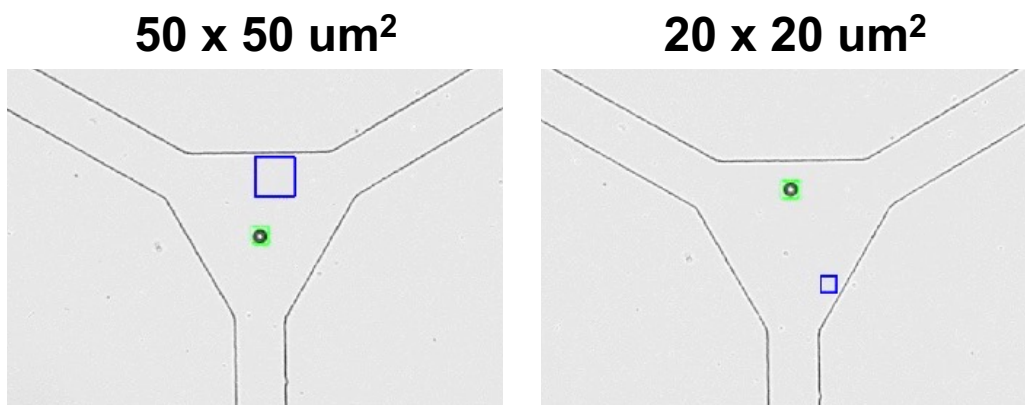


図 4.8 各ターゲットサイズ

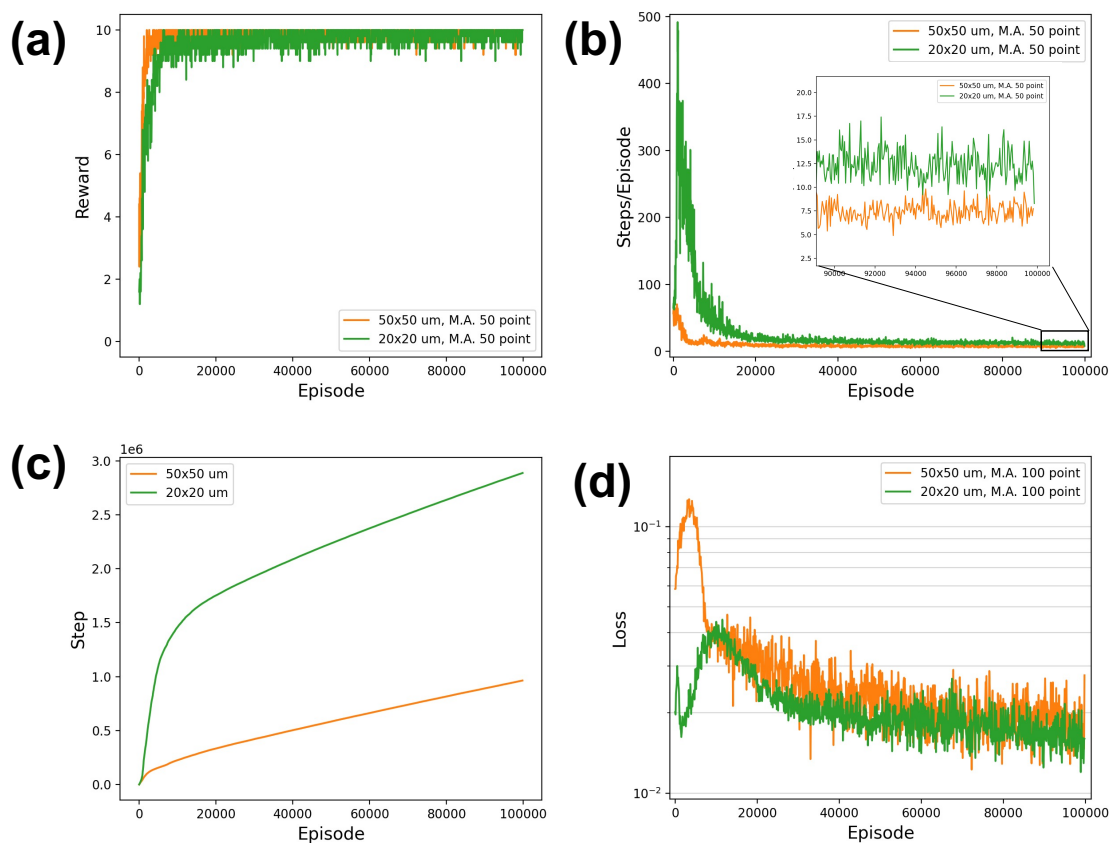


図 4.9 シミュレーションの結果 (a)報酬 (b)1 エピソードあたりのステップ数
(c) トータルステップ数 (d) 誤差の平均値

表 4.1 シミュレーションの成績

ターゲット	平均報酬	平均ステップ数 [step/episode]
50 x 50 um	9.95	7.68
20 x 20 um	9.83	11.88

表 4.2 実験の成績

ターゲット	平均報酬	平均ステップ数 [step/episode]
50 x 50 um	9.97	5.21
20 x 20 um	9.87	10.64

4.4 結言

本論文では、マイクロ流体システムを利用した新規マイクロマニピュレータのための液中微粒子の挙動の予測手法を提案した。本手法は、操作領域内の流れを CFD ソフトウェアで予測し、その流速データを微粒子の速度として 2D 物理エンジンで粒子の運動を予測する。実測値との比較から、直径 20 μm の粒子の平均移動量に対して、予測誤差が 17.6% となり、粒子の移動予測が可能であることが示された。

本シミュレータを用いて粒子の操作を実行した。まず、シミュレータでの学習が行われ、タスク成功率と操作効率の上昇から、2次元へ面における粒子操作が学習可能であることが確認された。その後、シミュレーションにおける学習結果を用いて実環境における操作を行った。50 \times 50 μm 、20 \times 20 μm のそれぞれのターゲットサイズで、タスク成功率は 98% 以上となった。ステップ数はシミュレーションと同じ程度であった。この結果から、マイクロバルブの強化学習による 2次元粒子操作技術の実証に成功したと考える。

第5章 深層学習を利用した粒子操作シミュレータと粒子の並列操作

5.1 緒言

本章では、流体中の粒子操作タスクをシミュレートするため、バルブ動作に対する粒子の挙動を予測する CFD を利用した手法をシミュレータに組み込んだ。シミュレータは一般的に、CFD や物理エンジンなどを含めて物理理論ベースで構築されるが、デバイスや操作対象の形状に合わせて一つ一つ数理モデルを作成することは、システムの応用範囲を制限する。この煩雑さを克服するため、深層学習ベースの手法を開発した。これは、バルブ動作に対する粒子の運動の関係をニューラルネットワークで予測する手法である。このネットワークは、粒子の位置とバルブの状態変化の情報を入力とし、バルブ動作後の粒子の位置を出力とするように訓練された。このネットワークは実際のデバイスからランダムサンプリングされた学習用データを用いた教師あり学習で実現される。

本章では、まず、2粒子の並列操作を実行できる操作自由度を持つデバイスを作製し、シミュレータ用のネットワークの学習用データの収集を行った。このデータを用いて粒子移動予測用のネットワークの学習を行った。ネットワークの予測精度を確認すると同時に、予測精度と学習データ数の関係を調べ、十分な精度を得るために必要な学習データ規模を評価した。その後、強化学習による粒子の操作の学習が行われた。構築したシミュレータ上でターゲットのサイズを変更し、システムの粒子操作精度の評価を行った。学習したモデルは実環境にデプロイされ、2次元平面上での粒子操作の実証が行われた。

5.2 理論、実験方法

5.2.1 デバイスの構造と強化学習の適用

図 5.1(a)にデバイスの構造を示す。デバイスは中央部の操作領域とそこにつながる流路、バルブで構成される。タスクの基本的な目的は、操作領域内の粒子の中心位置を任意のターゲットへ操作することである(図 5.1(b))。操作対象の粒子は、マイクロバルブの開閉で発生する流体の流れを利用して位置制御される。マイクロバルブは、Pneumatic layer の加圧と減圧によって開閉状態が切り替えられる(図 5.1(c))。操作領域につながる 6 本の流路に合計 7 つのバルブが配置された。2 つのバルブが配置された流路とその他の任意の 1 本の流路に注目したときに、一般的なマイクロポンプと同様にバルブが 3 つとなる。これは、任意の方向からシーケンシャルなバルブ動作で流体を出し入れすることができる最低限の構成である。 N 本の流路が交差した微小領域内における粒子操作の問題では、粒子数 P の操作において、2次元平面を操作するときの自由度は $2P$ で、これらの自由度を独立して操作するためには、独立した流量を生じる流路数が、 $N \geq 2P + 1$ を満たす必要がある[25]。よって、今回用いた 6 本の流路を持つデバイスでは 2 つの粒子の自由度を独立して操作できる。

図 5.1(d)は、本論文における粒子操作システム構築のスキームである。まず、実際のデバイスを用いてシミュレータのための学習データセットを収集する図 5.1(d, ①)。この学習データを用いて粒子の移動予測用のネットワークの学習を行う図 5.1(d, ②)。次のステップでは、シミュレータ上で深層強化学習を用いたエージェントの学習が実行される図 5.1(d, ③)。その後、シミュレータ上で十分に学習が行われたモデルが実環境にデプロイされ、実機の制御を行う図 5.1(d, ④)。

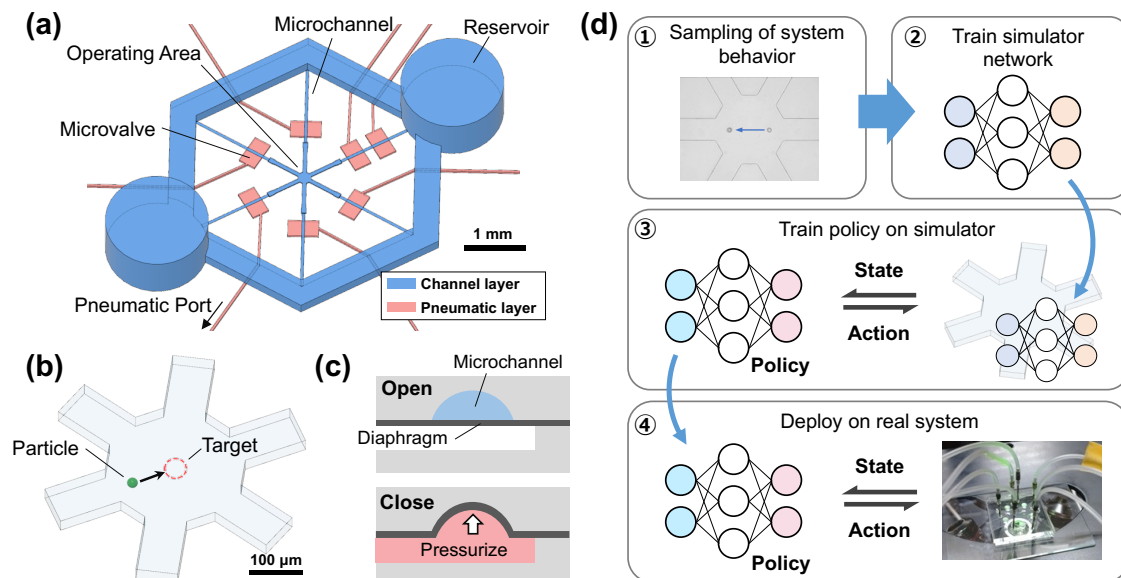


図 5.1 (a)流路構造の 3D イメージ (b)操作領域とタスクの概要 (c)マイクロバルブの断面イメージ (d)シミュレータの構築とエージェント学習のスキーム。

5.2.2 強化学習アルゴリズムの適用

本システムへの強化学習の適用のため、タスクの設定と特徴量の定義を行った。アルゴリズムなどは4章と同様である。ただし、ターゲット到達で得られる報酬は、 $r_t = 1$ とし、1回のタスクにおけるステップ数が500 stepを超えた場合も同様に次のエピソードに移行する。これらの場合も $r_t = 0$ である。

タスクの難易度が上がると、学習初期のエージェントが得る報酬が極端に少ないため、学習がほとんど進行しなくなる。これに対して、タスクを失敗した場合でも到達した状態に応じて報酬を与えることで、疎な報酬空間における学習効率を改善する、Hindsight experience replay [103]が提案されている。本論文では、この考え方を利用し、各遷移で収集される経験データのターゲット情報を次状態とし、報酬が与えられた場合の仮想的な経験を生成し、通常の replay buffer とは別の buffer に保存する。この buffer のキャパシティは replay buffer と同じとした。エージェントは学習時に replay buffer から経験データをランダムにサンプリングするが、ときにバッチサイズの10%を virtual buffer からサンプリングした経験を用いて学習を行った。

5.2.3 デバイスの作製

図5.2は、デバイスの流路構造である。Channel layer と Pneumatic layer のモールドは、フォトリソグラフィ[86]を用いて形成された。特に channel layer は、ポジ型とネガ型のフォトレジストを混在させる手法[104]を用いることで、バルブのシールのための半円形構造と矩形構造を一つのモールド上に構成した。モールドおよびデバイス製造は以下の手順で行われた。まず、ネガ型フォトレジスト(SU-8 3025, Nippon Kayaku Co., Ltd., Japan)をシリコンウェハにスピコートし、95°Cで30分ソフトベークした。このウェハをフォトマスクを使用せず、全面をUVライトに150 mJ/cm²の露光量で露光し、95°Cで5分間ポストベークした。このウェハ上にポジ型レジスト AZ P4903 (MicroChemicals GmbH)をスピコートし、110°Cのホットプレートで8分間ソフトベイクした。Rounded layer のフォトマスクを使用してUVライトに1000 mJ/cm²で露光した。その後、AZ Developer で現像された。このウェハをホットプレートで加熱し、AZ レイヤをリフローすることで半円形状の断面を得た。リフローは65°Cから190°Cまで10°C/hrで温度を上昇し、その後190°Cで固定され、合計15時間加熱した。このあと、リフロー後のウェハ上にネガ型フォトレジストで矩形流路部分を作製した。まず、Channel layer 1 がSU-8 3010で構成され、厚さは μm であった。ウェハを設計された流路が印刷されたフォトマスクを使用してUVライトに250 mJ/cm²で露光した。その後、ウェハを95°Cで5分間ポストベークした。同じ手順でSU-8 2100 (Nippon Kayaku Co., Ltd., Japan)を用いて channel layer 2 の構造を作成した。ただし、露光は400 mJ/cm²で行われた。Pneumatic layer のモールドはSU-8 3050のみで構成され、フォトマスクを介して200 mJ/cm²で露光で作製され、95.46 μm のレイヤー高さであった。

デバイスの厚さを調製するため、Channel layer は3 mm、Pneumatic layer は2 mmのシリ

コンゴムシートがスペーサとしてモールドに取り付けられた。PDMS (SYLGARD 184, The Dow Chemical Co., USA) は、base と curing agent が 10:1 の割合で混合され、モールドに流し込まれた。これを 75°C で 90 分加熱することで硬化させた。モールドからはく離された PDMS は、外形を切断し、チャンバーがパンチによって切断された。ダイアフラムは 10:0.7 の割合で混合された PDMS をシリコンウェハー上にスピコートされ、厚さを 17.0 μm とした。ダイアフラムと pneumatic layer に 90 Pa でプラズマを 25 秒間照射し、貼り合わせ、75° C のホットプレートで 15 分加熱することで接合した。このレイヤーと channel layer を同様の方法で接合した。

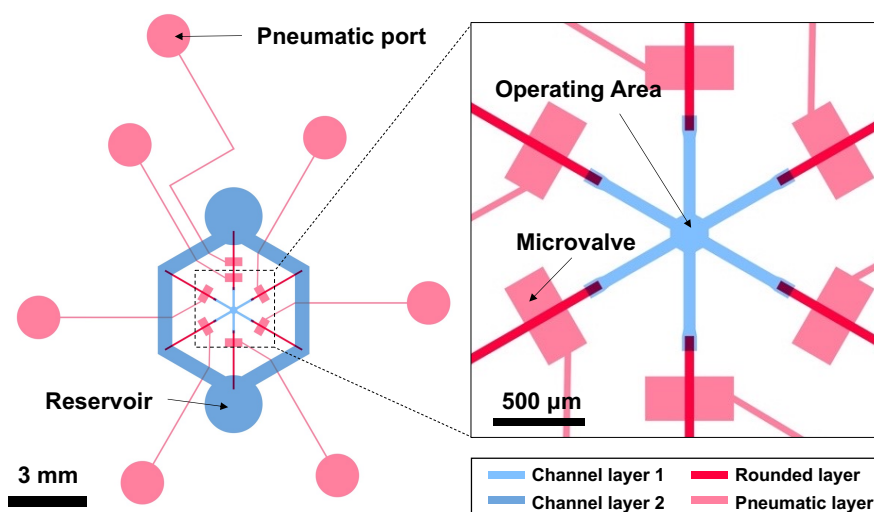


図 5.2 デバイス流路構造の 2D イメージ

5.2.4 学習システム

学習システムを図 5.3 に示す。システムは、ダイアフラムを駆動するための pneumatic control とソレノイドバルブ(M3QB110, CKD Corporation, Japan)、バルブを操作するための Arduino Microcontroller (Arduino Uno, Arduino S.R.L., Italy)、強化学習のプログラムを実行するためのコンピュータ、粒子を観察するための顕微鏡と CMOS カメラ(VC-320, Gazo Co., Ltd.)から構成される。デバイスは pneumatic chamber と空気の圧力をコントロールするためのダイアフラムポンプがソレノイドバルブを介して接続され、ソレノイドバルブは Arduino Microcontroller の制御信号によって駆動される。このとき pneumatic chamber に印加される圧力はすべての実験で 80 kPa とした。デバイスの流路内を、Milli-Q water に Tween20 (170-6531, Bio-Laboratories, Inc., USA) を 0.05%(v/v)で希釈したもので満たした。その後、デバイスの observation area に 20 μm のポリスチレン粒子(#18329, Polysciences Inc., USA)を配置する。顕微鏡に取り付けられたカメラから画像をコンピュータに取り込み、OpenCV 4.0 を用いて粒子の輪郭を抽出することで粒子位置をリアルタイムで測定する。

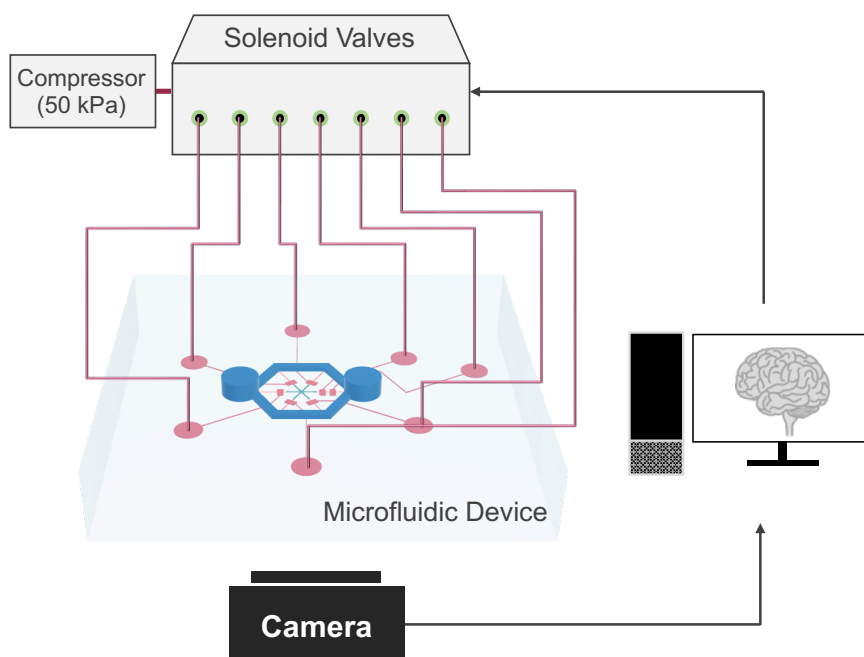


図 5.3 学習システム

5.2.5 シミュレータの構築

図 5.4 はシミュレータによる学習の概要を示す。エージェントは現在の状態から行動価値を戦略に従って出力し、シミュレータにアクション a を入力する。シミュレータは現在の粒子位置、バルブ状態と入力されたアクション a から、 a が実行されたあとの粒子の位置を出力する。エージェントはこの移動後の粒子の位置を用いて再度行動を選択する。これを繰り返して学習を行う。ここで、バルブ状態の遷移は、選択された行動のみで決まるとした。これは、実環境のシステムはポンプのダイヤフラムの動作は確率遷移せず、バルブはコントローラからの信号によって確実に想定した状態へ遷移することができるためである。

粒子位置を予測するネットワークの学習は、まず、ランダムにバルブを動かしたときの粒子の移動をサンプリングする。サンプリングの手順は以下の通りである。任意のバルブ状態 v_0 のときの粒子位置 p_0 を観測する。次に、ランダムに選択されたバルブ状態 v_1 へ遷移し、粒子位置 p_1 を観測する。その後、ランダムなバルブ遷移と粒子位置の観測を繰り返す。これにより、バルブ動作と粒子の挙動の関係を含む学習データセット

$$D = \langle [v_0, p_0, v_1, p_1], [v_1, p_1, v_2, p_2], \dots [v_{t-1}, p_{t-1}, v_t, p_t] \rangle \quad (4)$$

を生成する。次にこのデータセットにおける粒子初期位置 p_{t-1} と動作前後のバルブ開閉状態 v_{t-1}, v_t を入力とし、実測された移動先の座標 p_t を教師信号として教師あり学習でネットワークを訓練した。このネットワークは 256unit の隠れ層を 2 層持つ MLP で構成される。隠れ層の活性化関数は ReLU function、optimizer は Adam が用いられた。損失関数は Mean Absolute Error (MAE) が用いられた。

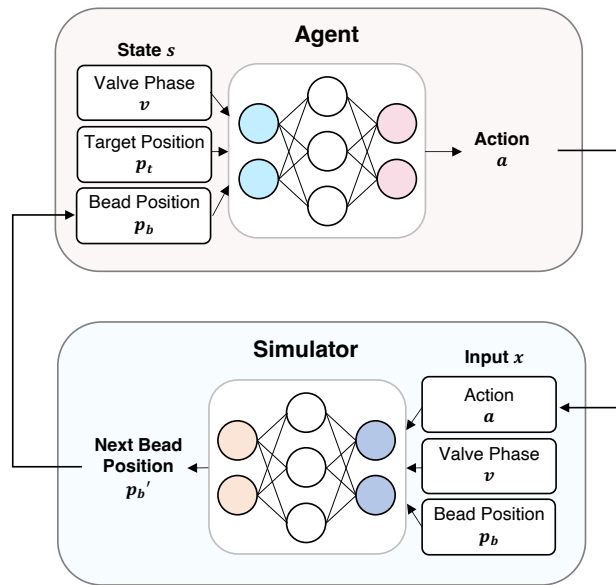


図 5.4 シミュレータの構造とこれを利用した強化学習の流れ

5.3 実験結果および考察

5.3.1 シミュレータの評価

シミュレータ構築のため、バルブ動作に対する粒子の移動予測を行うネットワークの学習を行った。まず、ランダムなバルブ動作で粒子の移動データの収集が行われ、約 55000 step のデータを取得した。このときサンプリングされた粒子の平均移動量は $19.82\ \mu\text{m}$ であった。訓練に必要なデータ数を見積もるため、データ数を変化させた場合の MAE を調べた(図 5.5(a))。学習データ数が増加するにつれてテストデータに対する予測誤差が低下していくことが示された。テスト誤差の低下が飽和するまでに 35000 データを要することが示された。使用したデバイスは7つのバルブを持ち、前後の開閉状態の組み合わせは 16384 通りである。確率的に各動作 2 回程度のサンプルがあれば領域全域に対して予測が可能となることが示唆された。50000 データで学習した場合のモデルの学習曲線を図 5.5(b)に示す。また、図 5.5(c)は、このモデルで予測された粒子の移動と実測された値の比較である。シミュレータ上の粒子初期位置をサンプリングした実環境における粒子遷移の初期位置に一致させ、バルブの動作後の粒子位置を算出し、比較した。図 5.5(b)より、このモデルの MAE は $1.47\ \mu\text{m}$ で、平均移動量に対して 7.44% であった。モデルは、壁などの境界条件の情報がモデルに与えられていないにも関わらず、定義された流路領域外への予測はほとんど見られなかった。これは、複雑な流路やピラーなどがある場合でもそれらを考慮した粒子の移動をデータから学習可能であることを示唆している。また、複数の粒子でのサンプリングに拡張することで粒子同士の相互作用を含めた予測ができる可能性がある。これらの性質は、変形を伴う粒子や細長い粒子など、粒子や流路形状の制約がほとんどなく、高速な予測が可能となるため、マイクロバルブの集積システム内の粒子の挙動予測手法として利用が期待できる。

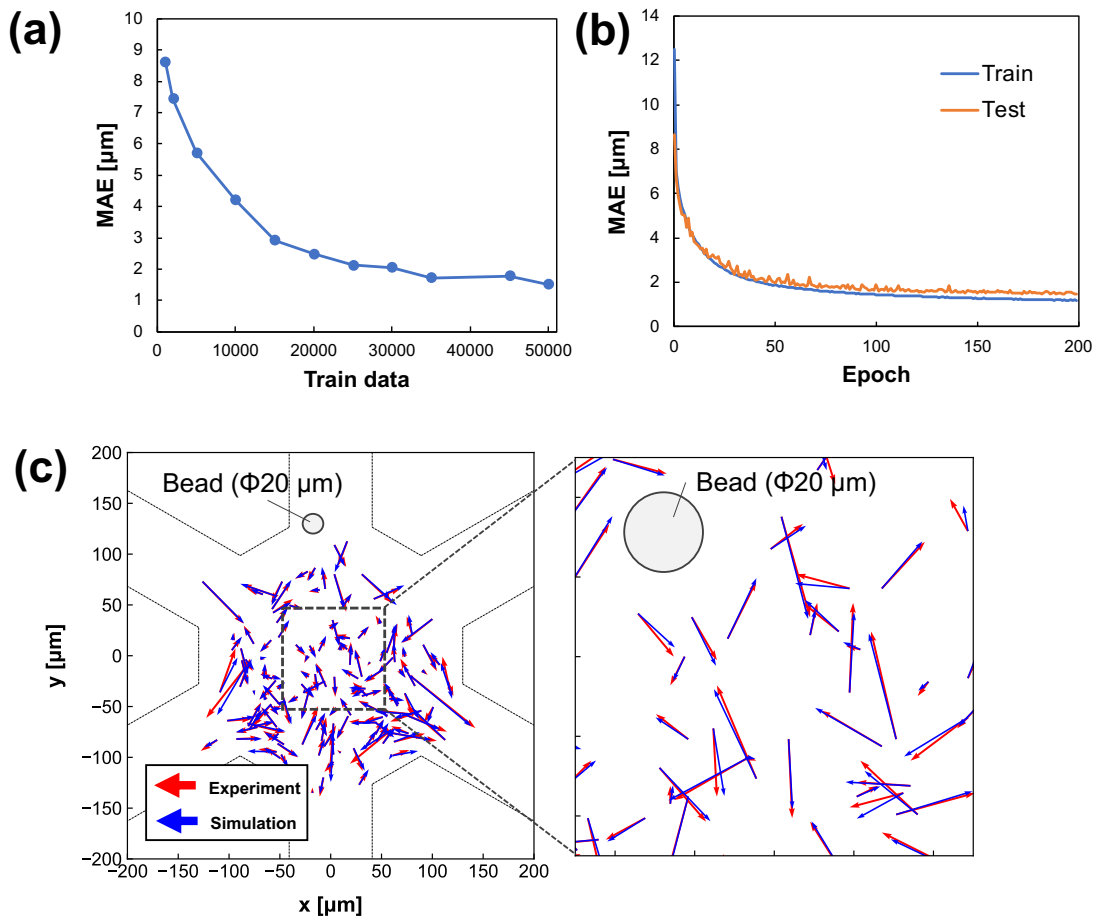


図 5.5 学習データ数と予測誤差の関係(a) データ数に対する MAE の変化。すべての試行で 5000 step がテスト用データとして割り当てられた。(b) MAE の推移。256 unit の hidden layer が 2 層のモデル, $\alpha = 0.0001$, バッチサイズ 32。(c)予測結果と実測値の比較。使用モデルは(b)と同様。

5.3.2 単一粒子操作

実環境での実験に先立ち、シミュレータを用いた 1 粒子の操作の学習が行われた。図 5.6(a, b)にシミュレータによる学習結果を示す。図 5.6(a,b)より、ターゲットエリア直径を $10\ \mu\text{m}$ から縮小していくと、直径 $2.5\ \mu\text{m}$ まではタスク成功率が 90%を超え、ステップ数も効率化されることが確認された。 $1\ \mu\text{m}$ ではタスク成功率が 40%程度となり、それ以上は上昇しなかった。step 数は、ターゲット付近での微調整の難易度が上がるため、ターゲットエリアの縮小に比例して増加する傾向がある。さらに、 $1\ \mu\text{m}$ のタスクでは、カメラの解像度は 1 px あたり約 $0.9\ \mu\text{m}$ であるため、画像の摂動により位置調整の難易度が上昇したことが原因であると考えられる。

ここで、経験の改変の効果を確認するため、この手法の有無によるシミュレーション結

果の比較を行った。図 5.7 に経験の改変の有無によるシミュレーションの結果を示す。基本的に、どのターゲットエリアのタスクにおいても報酬の立ち上がりやステップ数の減少が早い段階から確認された。さらに、10, 5 μm では最終的には再ラベリングを使用しない場合でも同等のパフォーマンスが得られたが、2.5 μm では経験の改変なしではタスクを学習することができなかった。これは、タスクの難易度が上がり、学習が不足している状態では報酬にたどり着くことができず、学習が進行しないことによるものである。このシミュレーション結果から、ターゲットサイズ縮小により報酬が疎な場合に経験の改変による学習が有効であることが示された。

図 5.6(c,d)はシミュレーションと実環境にデプロイしたときの結果である。さらに、図 5.8 は操作中の様子である。バルブの動作で発生する流れを利用して指定されたターゲットエリアに粒子を運ぶ動作が確認された。実環境においては、直径 2.5 μm のタスクにおける最終的な平均ステップ数は約 60 steps/episode となり、バルブ動作スピードは約 3 Hz 程度であるため、目的のエリアへの移動に平均 20 秒程度で粒子を配置することが可能である。実環境にシミュレーション結果をデプロイすると、シミュレーションで獲得した動作によって、デプロイした直後から良い成績が得られる。基本的にシミュレータと実環境にはギャップが存在し、完全に同じ環境での学習は困難であるため、一般に、このギャップによって実環境においてシミュレーションの成績を再現することが妨げられる。しかし、提案手法では、デプロイしたエージェントは、シミュレーションで最終的に到達した結果には及ばないものの、それに近い成績を達成し、十分な再現性が確保されている。この事実、シミュレータによるパラメータ調整や学習効率の評価の精度が確保されていることを示している。

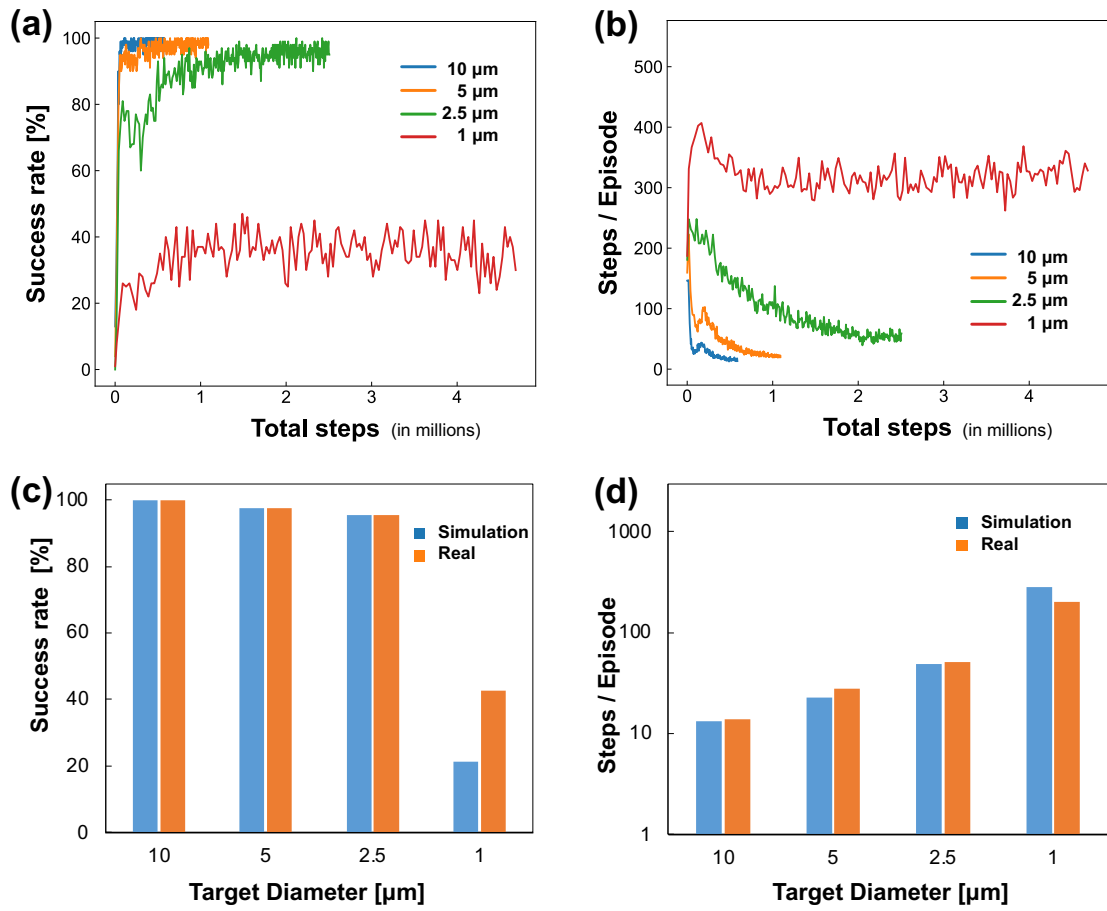


図 5.6 単一粒子操作のシミュレーション及び実環境におけるテストの結果。(a)シミュレーションにおけるタスク成功率と(b)1 エピソードあたりのステップ数。 $\alpha = 0.0001$ 、 $\gamma = 0.98$ とした。各プロットは 100 エピソードの平均値である。(c) 各環境におけるタスク成功率 (d) 1 エピソードあたりのステップ数 (200 episode の平均)

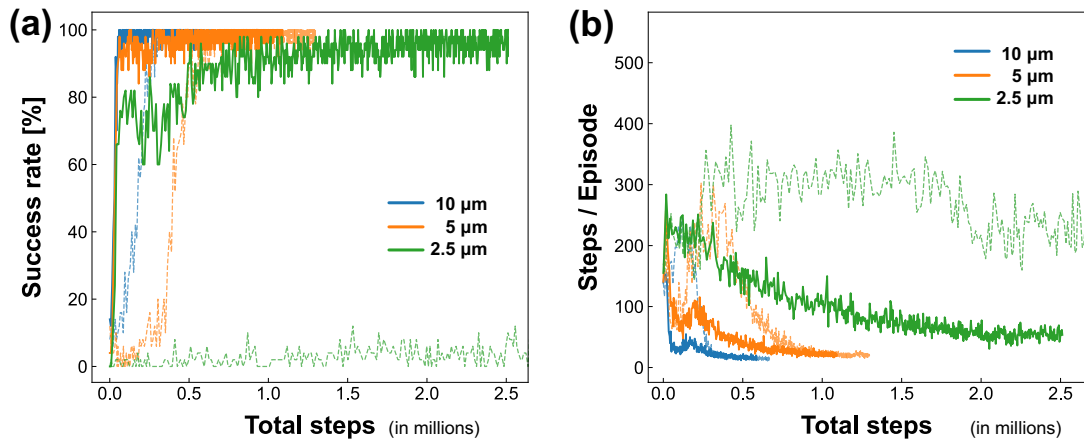


図 5.7 1 粒子操作の結果(a)タスク成功率と(b)1 エピソードあたりのステップ数。各プロットは 50 エピソードの平均値である。破線は各ターゲットエリアにおける報酬の改変を行わなかった場合の結果。 $\alpha = 0.0001$ 、 $\gamma = 0.98$ 。

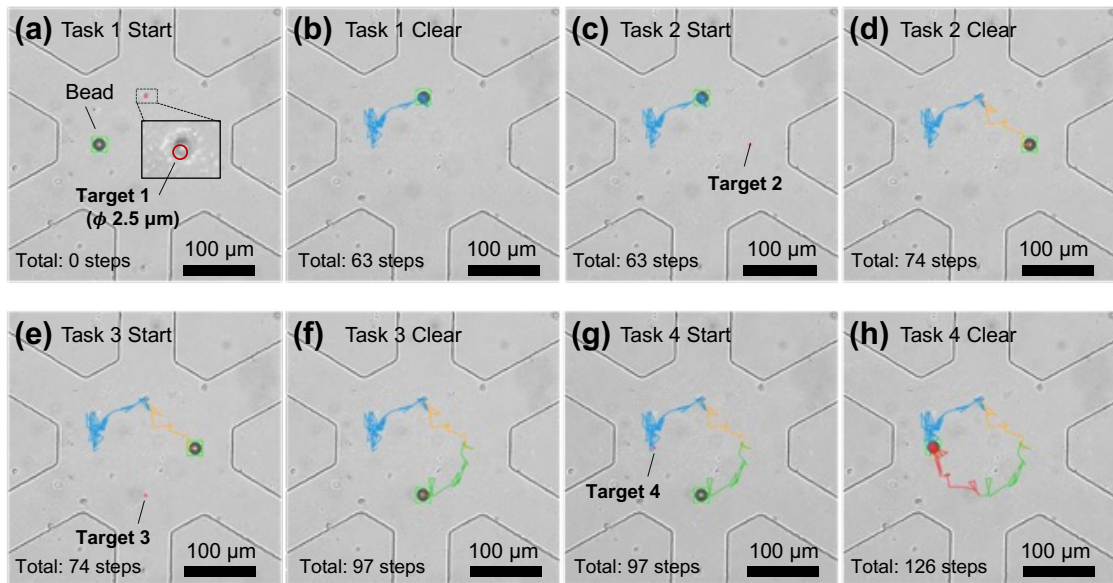


図 5.8 実環境における連続した 4 回の試行の写真。ターゲット $\phi 2.5 \mu\text{m}$ 。動作速度 3 Hz。図中のラインは粒子の軌道を示す。(a,b) Task 1 (c,d) Task 2 (e,f) Task 3 (g,h) Task 4

5.3.3 粒子 2 個並列操作

次のタスクとして、領域内の 2 つの粒子の同時操作を学習した。ターゲットエリア同士はオーバーラップしないこととし、シミュレータのネットワークは粒子 1 個の場合と同じものを継続して使用した。移動予測は各粒子に対して独立して行われるため、シミュレータ内の予測では粒子同士のオーバーラップなどは許容されている。その他の学習設定は 1 粒子の場合と同様である。図 5.9(a, b) にシミュレータの学習結果を示す。20 μm のターゲットの場合、最終 1000 episodes のタスクの成功率が 98.8%、1 エピソードあたりのステップ数は平均 32.74 steps となった。ターゲットを縮小すると、10 μm ではタスク成功率が 70%程度、5 μm では最大約 15%となった。ターゲットが縮小された場合でも、経験の改変操作によって異なるタスクにおける報酬が得られる動作を学習できるが、ターゲット周辺にたどり着くまでの動作を学習することができないことが原因と考えられる。この結果を受け、20 μm のターゲットで学習したモデルを 10 μm のターゲット設定で追加学習を行った(図 5.9, 10 μm (Pre-trained))。この場合、初期段階から高いタスク成功率を示し、学習の進行に伴いステップ数も減少した。以上の結果より、シミュレータ上での複数粒子の並列操作が可能であることが確認された。

ここでも、経験の改変の効果が確認された。図 5.10 に経験の改変の有無による 2 粒子の並列操作タスクにおけるシミュレーションの結果を示す。2 粒子の場合、経験の改変なしでは 20 μm のタスクを学習することができなかった。一方で、再ラベリングを行った場合ではタスク成功率の上昇と動作の効率が確認された。よって、2 粒子の並列操作の場合でもこの手法の効果が確認された。

実環境におけるパフォーマンスを実証するため、シミュレータで学習を行ったモデルを実環境へ適用した。図 5.9(c, d) に実環境とシミュレータのパフォーマンスの比較結果を示す。1 粒子の場合と同じように、シミュレータでは初期状態を実環境のデータに一致させてタスク成績を算出した。図 5.11 は操作中の様子である。実環境においても粒子の並列操作に成功していることを確認した。ステップ数はシミュレーションに比べて効率がわずかに低下するものの、2 つの環境で大きなギャップは見られなかった。これにより、粒子数を増やした場合でもシミュレータを用いた学習ができることが示された。また、粒子間同士の相互作用を考慮していない移動予測ネットワークを用いても同時操作を学習できることもわかった。現在のシステムは 3 Hz で動作しているため、任意の位置へ、20 μm では約 20 秒、10 μm では約 50 秒で操作可能である。領域内にピラーや堰などを設け、それらのツールを利用した操作を学習させることで、流路数と操作可能な粒子数の間にある流体力学的な制約にとらわれない、効率的なシステムの実現に期待できる。

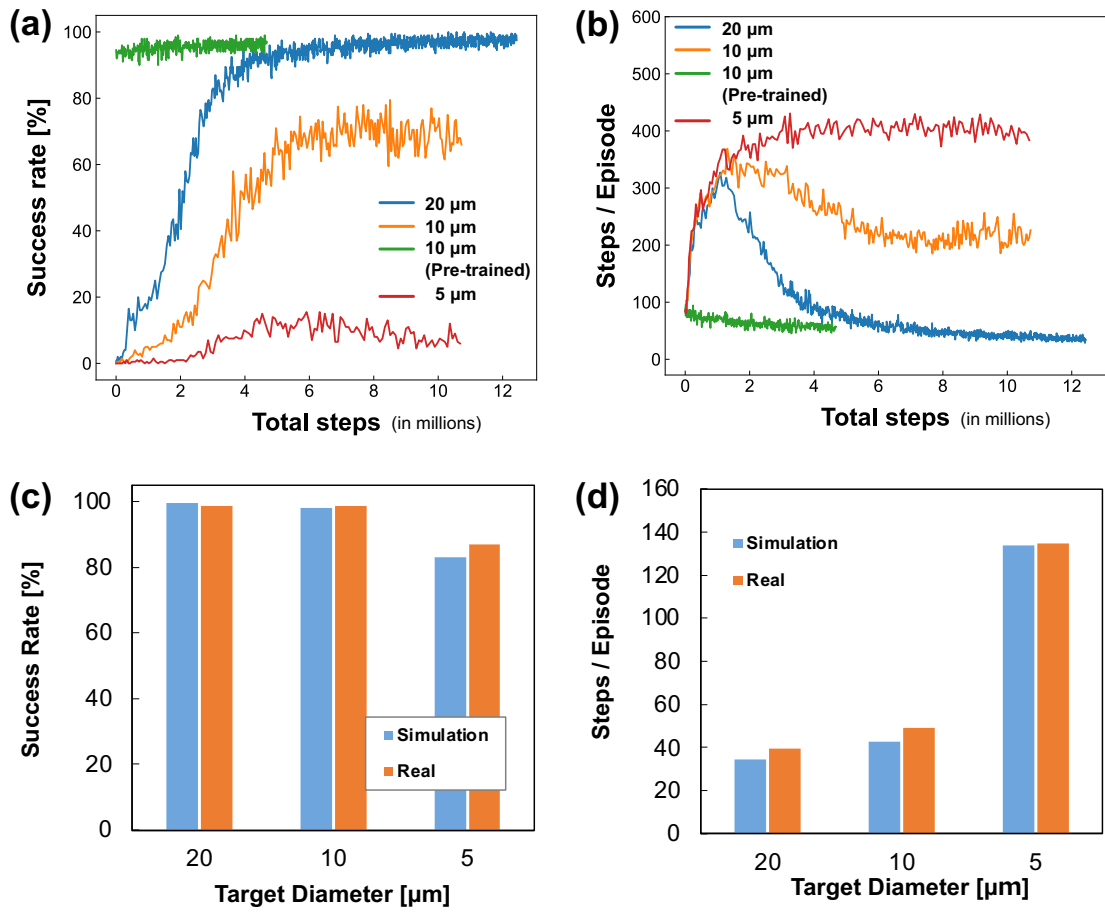


図 5.9 2 粒子並列操作のシミュレーション及び実環境におけるテストの結果。(a)シミュレーションにおけるタスク成功率と(b)1 エピソードあたりのステップ数。各プロットは 200 エピソードの平均値である。Pre-trained のプロットは 20 μm での学習済みモデルを使用した場合の結果。このとき $\alpha = 0.000001$ 、 $\gamma = 0.98$ とし、それ以外は $\alpha = 0.00001$ とした。(c) 各環境におけるタスク成功率と(d)1 エピソードあたりのステップ数 (200 episode の平均、10 μm のプロットは(a,b)における Pre-trained のモデル)。

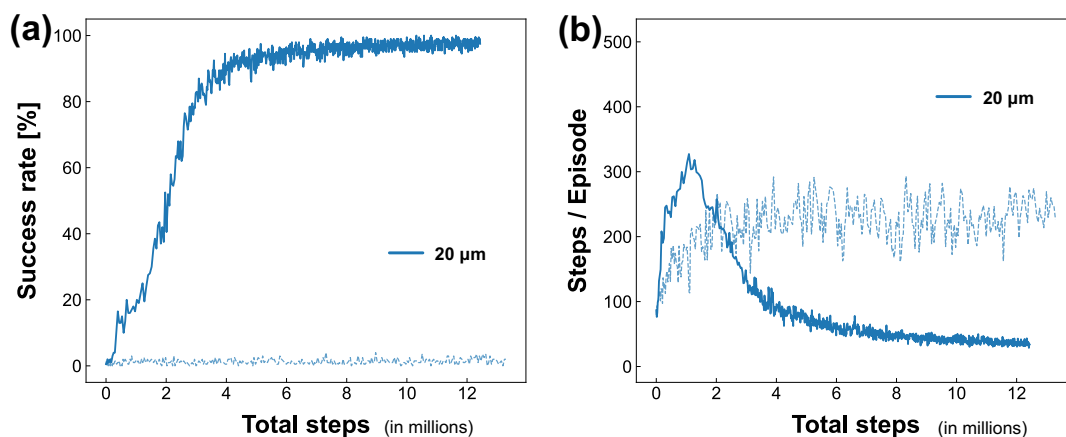


図 5.10 1 粒子操作のシミュレーション結果。(a)タスク成功率と (b) 1 エピソードあたりのステップ数。各プロットは 200 エピソードの平均値である。破線は報酬の改変を行わなかった場合の結果。 $\alpha = 0.00001$ 、 $\gamma = 0.98$ 。

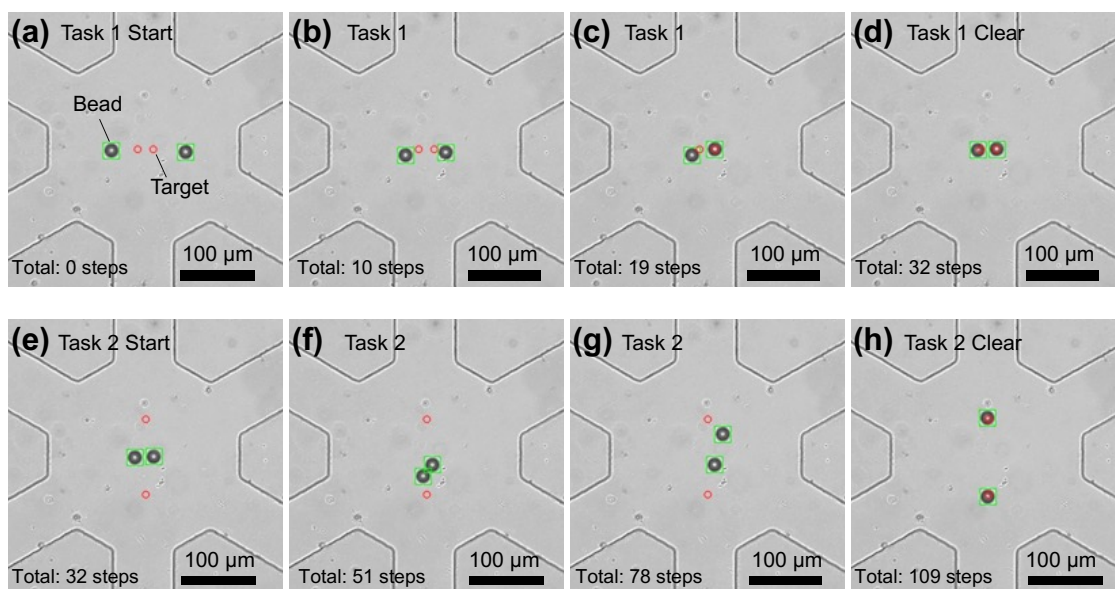


図 5.11 実環境における連続した 2 回の試行の写真。ターゲット $\phi 2.5 \mu\text{m}$ 。動作速度 3 Hz。
(a-d) Task 1 (e-h) Task 2

5.4 結言

本章では、マイクロバルブを用いた 2 次元平面での新しい粒子操作技術について報告した。このシステムは、オンチップマイクロバルブの動作で発生する流れを用いて粒子を操作する。適切なバルブシーケンスは強化学習を利用することで自律的に決定された。エージェントの効率的な学習のため、シミュレータが用いられた。タスクをシミュレートするために、微小流路内の粒子の運動を予測する必要がある。そこで、我々は深層学習を利用してその運動を予測した。バルブをランダムに動かし領域内の粒子の運動をサンプリングし、バルブの動作と粒子の位置から動作後の粒子位置を予測するニューラルネットワークを教師あり学習で訓練した。予測誤差が平均 7.44% となり粒子位置の予測が可能であることが示された。このシミュレータを用いてエージェントの学習を行った。まず、1 粒子の操作を行い、任意の位置に操作可能であることが示された。その後、2 粒子の同時操作を行い、2 つの粒子を同時に操作してターゲットに運ぶことができることを確認した。

本手法は、マイクロバルブと微小流路構造のみで構成され、マイクロ流体マルチプレクサを利用することで大規模並列化することが可能となる。また、強化学習を利用したことでデバイスの流路構造や粒子の物理的特性に制約がない。さらに、シミュレータにおいてもこれらの制約はない。今後、ピラーや堰などの構造物を利用した行動を獲得させることで、流路の本数と操作可能粒子数の制約を受けることなく操作できるかもしれない。したがって、細胞ネットワークの解析プラットフォームや機能性粒子相互作用の基礎的特性を調査するプラットフォームとしての応用に期待できる。

第6章 総括

本論文では、マイクロ流体システムの強化学習による制御技術と、これを応用した安価で生体適合性のある微粒子操作技術を提案した。

第1章では、マイクロ流体システムの強化学習による制御技術の確立のため、マイクロ流体工学でよく用いられるオンチップアクチュエータのマイクロバルブを制御対象として強化学習を適用した。ここでは、3つのマイクロバルブを組み合わせたマイクロペリスタルティックポンプと呼ばれる構造を用いた。この機構を搭載したマイクロ流体デバイスを作製し、バルブの動作機構と、カメラなどの観察機構、情報処理用のコンピュータを統合した学習システムを構築した。本章では、オンチップアクチュエータへの強化学習の適用の原理を実証するため、マイクロポンプの送液動作の獲得を目標とした。これに合わせて、マイクロポンプの動作にマルコフ性を仮定し、MDPに基づく定式化を行った。バルブの動作で発生する、順方向への流れを報酬とすることで、システムの目的を流量の最大化とした。その後、実際のデバイスにおける流体の移動量のサンプリングを行い、報酬関数のモデリングが行われた。続いて、MDPにおける価値関数の計算のため、動的計画法による最適化が行われた。コンピュータ上で学習が行われ、割引率の最適化により、従来提案されていたシーケンスに比べて大幅に流量が向上する動作を学習できることが示された。また、得られた動作シーケンスは本実験で使用したシステムの動作特性に応じて最適化されたものであることが確認された。

動的計画法を適用する場合、報酬関数が計算可能である必要がある。しかし、実際の環境では、環境が未知の場合が多く、報酬関数などの数理モデルを構築することは難しい。そのため、環境内を探索し試行錯誤で最適な行動を学習するモデルフリー型アルゴリズムのQ学習の適用を試みた。実環境での学習に加え、MDPの適用時に作成した報酬関数を組み込んだ仮想環境を用いてコンピュータ上でも同様の学習が行われた。いずれの環境においても動的計画法で獲得されたものと同様のシーケンスに収束し、正常に学習が行われたことを確認した。報酬を目標流量からの偏差に変更し、順方向への流量の調整を目的として、より高度な動作の学習を行った。いくつかの流量の設定において、所望の平均流量を発生する動作を獲得できることを確認した。さらに、学習中に空気圧制御系の電磁弁を変更し環境を変化させた場合でも、パフォーマンスを維持できることを示した。

第2章では、強化学習によるマイクロバルブの制御技術を拡張することで粒子操作技術を実証した。直線上の流路内の粒子を指定した位置へ運ぶタスクとして特徴量が定義された。粒子がターゲットエリアに到達したときに報酬を与え、粒子を目的位置に運ぶような動作の学習を目的とした。学習アルゴリズムにはQ学習が用いられ、シミュレータ、実環境のそれぞれの環境で学習が行われた。実環境における実験によって、ランダムに動作させた場合に比べて効率よく粒子を操作できることを確認した。また、シミュレータを用いた事前学習を行うことで、実環境における学習を効率化できることが確認された。

第3章では、マイクロバルブの作製手法として、EVA ホットメルト接着剤と VAM を用いたリフロープロセスを報告した。この手法は、溶融した EVA グルーを、マイクロ流路構造を用いて成形することで、加熱によって断面形状を制御可能なマイクロ流体デバイス用のモールド作製技術である。この手法は、従来のポジ型フォトレジストでは困難な、数百 μm スケールの半円形流路断面を作製することができる。まず、シリコンウェハ上への直線流路のパターニングを行い、加熱時間によって断面形状を CV が 5 % 以内で制御可能であることを示した。その後、曲線流路を用いたパターニング能力の評価を行った。直角に曲がる流路の曲線部におけるリフロー前後の流路幅の変化を調べ、複雑な流路構造に対応できることを確認した。その後、バルブ作製と、マイクロペリスタルティックポンプを作製した。蛍光観察やポンプ流量の測定を行い、本手法がマイクロ流体デバイスの作製に適用可能であることを確認した。

第4章では、強化学習を利用した新規マイクロマニピュレータのための液中微粒子挙動の予測手法を提案した。これは、操作領域内の流れを CFD ソフトウェアで予測し、その流速データを微粒子の速度として 2D 物理エンジンで粒子の運動を予測する。実測値との比較から、直径 20 μm の粒子の平均移動量に対して、予測誤差が 17.6% であった。これによって、液中粒子の移動予測能力が実証された。その後、このシミュレータを用いて粒子の操作を実行した。まず、シミュレータでの学習により、タスク成功率と操作効率の上昇が確認され、2次元平面での粒子操作が可能であることが示された。この結果から、シミュレーションにおける学習結果を用いて実環境での粒子操作を実行した。ターゲットサイズが 50×50 μm および 20×20 μm で、タスク成功率は 98% 以上となった。また、初期位置からターゲットまでのステップ数はシミュレーションと同程度であった。以上の結果からマイクロバルブの強化学習による 2次元粒子操作技術が実証された。

第5章では、深層学習を利用したシミュレータを構築し、複数粒子の同時操作を実現した。システムの動作原理は第2章、第4章と同じであるが、フォトリソを用いたデバイス作製により並列化が可能である。学習に先立ち、シミュレータの構築を行った。粒子の挙動を予測するため、深層学習が用いられた。実際のデバイスでバルブをランダムに動かし、粒子の挙動をサンプリングした。このデータセットを用いて、バルブ動作と現在の粒子位置から次の粒子位置を予測するニューラルネットワークを訓練した。このネットワークは、粒子の平均移動量に対する予測誤差が 7.44% となり、粒子移動予測が可能であることが示された。このシミュレータを用いて強化学習モデルの学習を行った。はじめに、1粒子の操作を行い、指定された任意のターゲット位置へ粒子を操作できることを確認した。その後、2つの粒子の同時操作を実行し、それらを同時に操作して目標位置に移動させる動作を学習したことを確認した。これらの結果は、シミュレータでの学習を行った後、実環境に適用された。実環境に置いてもシミュレータで学習した戦略によってパフォーマンスが維持できることが示された。操作精度は、1粒子の場合は 2.5 μm 、2粒子の同

時操作の場合は $10\ \mu\text{m}$ となった。以上の結果より、2次元平面における粒子操作技術が確立された。

本研究では、オンチップマイクロ流体アクチュエータの強化学習による制御を用いた粒子操作技術の構築を目標とし、制御手法とシステムの開発を行った。本研究で確立した技術は、低コストで集積可能な微小粒子操作技術として、細胞生物学を始めとした分野での利用に期待できる。また、粒子操作だけでなく、強化学習による自律的な動作の獲得により、目標を設定するのみで複雑な動作が可能となることも同時に示した。これは、タスクに対して専用にデバイスを作製する必要があり、汎用性に乏しいという問題があった、マイクロ流体システムの問題の解決に貢献するものである。

参考文献

- [1] K. Cheng, J. Guo, Y. Fu, J. Guo, Active microparticle manipulation: Recent advances. *Sens. Actuators A Phys.*, 322, 112616 (2021).
- [2] J. Nilsson, M. Evander, B. Hammarström, T. Laurell, Review of cell and particle trapping in microfluidic systems. *Anal. Chim. Acta.* 649, 141–157 (2009).
- [3] D. A. Ateya, J. S. Erickson, P. B. Howell, L. R. Hilliard, J. P. Golden, F. S. Ligler, The good, the bad, and the tiny: a review of microflow cytometry. *Anal. Bioanal. Chem.* 391, 1485–1498 (2008).
- [4] N. Pamme, Continuous flow separations in microfluidic devices. *Lab Chip.* 7, 1644–1659 (2007).
- [5] P. Y. Chiou, A. T. Ohta, M. C. Wu, Massively parallel manipulation of single cells and microparticles using optical images. *Nature.* 436, 370–372 (2005).
- [6] D. Spetzler, J. York, C. Dobbin, J. Martin, R. Ishmukhametov, L. Day, J. Yu, H. Kang, K. Porter, T. Hornung, W. D. Frasch, Recent developments of bio-molecular motors as on-chip devices using single molecule techniques. *Lab Chip.* 7, 1633–1643 (2007).
- [7] H. Li, D. Zhou, H. Browne, Evidence for resonance optical trapping of individual fluorophore-labeled antibodies using single molecule fluorescence spectroscopy. *J. Am. Chem. Soc.* 128, 5711–5717 (2006).
- [8] C. Gosse, V. Croquette, Magnetic tweezers: micromanipulation and force measurement at the molecular level. *Biophys. J.* 82, 3314–3329 (2002).
- [9] D. R. Gossett, W. M. Weaver, A. J. Mach, S. C. Hur, H. T. K. Tse, W. Lee, H. Amini, D. Di Carlo, Label-free cell separation and sorting in microfluidic systems. *Anal. Bioanal. Chem.* 397, 3249–3267 (2010).
- [10] L. Kang, B. G. Chung, R. Langer, A. Khademhosseini, Microfluidics for drug discovery and development: from target selection to product lifecycle management. *Drug Discov. Today.* 13, 1–13 (2008).
- [11] P. S. Dittrich, A. Manz, Lab-on-a-chip: microfluidics in drug discovery. *Nat. Rev. Drug Discov.* 5, 210–218 (2006).
- [12] Y. Yang, Y. Ren, M. Chen, Y. Arita, C. Rosales-Guzmán, Optical trapping with structured light: a review. *Advanced Photonics.* 3, 034001 (2021).
- [13] F. M. Fazal, S. M. Block, Optical tweezers study life under tension. *Nat. Photonics.* 5, 318–321 (2011).
- [14] A. Ashkin, J. M. Dziedzic, Optical trapping and manipulation of viruses and bacteria. *Science.* 235, 1517–1520 (1987).
- [15] M. B. Rasmussen, L. B. Oddershede, H. Siegumfeldt, Optical tweezers cause physiological damage to *Escherichia coli* and *Listeria* bacteria. *Appl. Environ. Microbiol.* 74, 2441–2446 (2008).
- [16] X. Wang, C. Ho, Y. Tsatskis, J. Law, Z. Zhang, M. Zhu, C. Dai, F. Wang, M. Tan, S. Hopyan, H. McNeill, Y. Sun, Intracellular manipulation and measurement with multipole magnetic tweezers. *Sci*

Robot. 4 (2019).

[17] M. D. Armani, S. V. Chaudhary, R. Probst, B. Shapiro, Using feedback control of microflows to independently steer multiple particles. *J. Microelectromech. Syst.* 15, 945–956 (2006).

[18] F. Akkoyun, S. Gucluer, A. Ozcelik, Potential of the acoustic micromanipulation technologies for biomedical research. *Biomicrofluidics.* 15, 061301 (2021).

[19] P. Zhang, C. Chen, F. Guo, J. Philippe, Y. Gu, Z. Tian, H. Bachman, L. Ren, S. Yang, Z. Zhong, P.-H. Huang, N. Katsanis, K. Chakrabarty, T. J. Huang, Contactless, programmable acoustofluidic manipulation of objects on water. *Lab Chip.* 19, 3397–3404 (2019).

[20] F. Guo, Z. Mao, Y. Chen, Z. Xie, J. P. Lata, P. Li, L. Ren, J. Liu, J. Yang, M. Dao, S. Suresh, T. J. Huang, Three-dimensional manipulation of single cells using surface acoustic waves. *Proc. Natl. Acad. Sci.* 113, 1522–1527 (2016).

[21] A. Ozcelik, J. Rufo, F. Guo, Y. Gu, P. Li, J. Lata, T. J. Huang, Acoustic tweezers for the life sciences. *Nat. Methods.* 15, 1021–1028 (2018).

[22] Z. Tian, S. Yang, P.-H. Huang, Z. Wang, P. Zhang, Y. Gu, H. Bachman, C. Chen, M. Wu, Y. Xie, T. J. Huang, Wave number-spiral acoustic tweezers for dynamic and reconfigurable manipulation of particles and cells. *Sci Adv.* 5, eaau6062 (2019).

[23] B. R. Lutz, J. Chen, D. T. Schwartz, Hydrodynamic tweezers: 1. Noncontact trapping of single cells using steady streaming microeddies. *Anal. Chem.* 78, 5429–5435 (2006).

[24] A. Shenoy, C. V. Rao, C. M. Schroeder, Stokes trap for multiplexed particle manipulation and assembly using fluidics. *Proc. Natl. Acad. Sci.* 113, 3976–3981 (2016).

[25] T. M. Schneider, S. Mandre, M. P. Brenner, Algorithm for a microfluidic assembly line. *Phys. Rev. Lett.* 106, 094503 (2011).

[26] N. Convery, N. Gadegaard, 30 Years of Microfluidics. *Micro Nano Eng.* 2, 76-91 (2019).

[27] R. Dong, Y. Liu, L. Mou, J. Deng, X. Jiang, Microfluidics-Based Biomaterials and Biodevices. *Adv. Mater.* 31, 1805033 (2019).

[28] G. M. Whitesides, The origins and the future of microfluidics, *Nature*, 442, 368-373 (2006).

[29] J. R. Lake, K. C. Heyde, W. C. Ruder, Low-cost feedback-controlled syringe pressure pumps for microfluidics applications. *PLoS One.* 12, e0175089 (2017).

[30] A. K. Fajrial, A. Vega, G. Shakya, X. Ding, A frugal microfluidic pump. *Lab Chip.* 21, 4772–4778 (2021).

[31] W. Gao, M. Liu, S. Chen, C. Zhang, Y. Zhao, Droplet microfluidics with gravity-driven overflow system. *Chem. Eng. J.* 362, 169–175 (2019).

[32] J.-Y. Qian, C.-W. Hou, X.-J. Li, Z.-J. Jin, Actuation Mechanism of Microvalves: A Review. *Micromachines.* 11, 172 (2020).

[33] H.-P. Chou, M. A. Unger, S. R. Quake, A Microfabricated Rotary Pump. *Biomedical Microdevices,* 3, 4, 323-330 (2001).

- [34] K. Leung, H. Zahn, T. Leaver, K. M. Konwar, N. W. Hanson, A. P. Page, C.-C. Lo, P. S. Chain, S. J. Hallam, C. L. Hansen, A programmable droplet-based microfluidic device applied to multiparameter analysis of single microbes and microbial communities. *Proc. Natl. Acad. Sci.* 109, 20, 7665-7670 (2012).
- [35] L. M. Fidalgo, S. J. Maerkl, A software-programmable microfluidic device for automated biology. *Lab Chip*. 11, 1612–1619 (2011).
- [36] T. Thorsen, S. J. Maerkl, S. R. Quake, Microfluidic large-scale integration. *Science*. 298, 580–584 (2002).
- [37] T. Zhang, Z.-Y. Hong, S.-Y. Tang, W. Li, D. W. Inglis, Y. Hosokawa, Y. Yalikun, M. Li, Focusing of sub-micrometer particles in microfluidic devices. *Lab Chip*. 20, 35–53 (2020).
- [38] S. Dekker, W. Buesink, M. Blom, M. Alessio, N. Verplanck, M. Hihoud, C. Dehan, W. César, A. Le Nel, A. van den Berg, M. Odijk, Standardized and modular microfluidic platform for fast Lab on Chip system development. *Sens. Actuators B Chem.* 272, 468–478 (2018).
- [39] K. A. Shaikh, K. S. Ryu, E. D. Goluch, J.-M. Nam, J. Liu, C. S. Thaxton, T. N. Chiesl, A. E. Barron, Y. Lu, C. A. Mirkin, C. Liu, A modular microfluidic architecture for integrated biochemical analysis. *Proc. Natl. Acad. Sci.* 102, 9745–9750 (2005).
- [40] H. Wang, Y.-L. Zhang, D.-D. Han, W. Wang, H.-B. Sun, Laser fabrication of modular superhydrophobic chips for reconfigurable assembly and self-propelled droplet manipulation. *Photonix*. 2, 1–13 (2021).
- [41] E. A. Galan, H. Zhao, X. Wang, Q. Dai, W. T. S. Huck, S. Ma, Intelligent microfluidics: The convergence of machine learning and microfluidics in materials science and biomedicine. *Matter*. 3, 6, 1893-1922 (2020).
- [42] J. Zheng, T. Cole, Y. Zhang, J. Kim, S.-Y. Tang, Exploiting machine learning for bestowing intelligence to microfluidics. 194, 15, 113666 (2021).
- [43] V. Rizzuto, A. Mencattini, B. A.-Gonzalez, D. D. Giuseppe, E. Martinelli, D. B.-Pastor, M. d. M. M.-Pereira, M. J. L.-Martinez, J. Samitier, Combining microfluidics with machine learning algorithms for RBC classification in rare hereditary hemolytic anemia. *Sci. Rep.* 11, 13553 (2021).
- [44] S. Ota, R. Horisaki, M. Ugawa, I. Sato, K. Hashimoto, R. Kamesawa, K. Setoyama, S. Yamaguchi, K. Fujiu, K. Waki, H. Noji, Ghost cytometry. *Science*, 360, 6394, 1246-1251 (2018).
- [45] N. Nitta, T. Sugimura, A. Isozaki, H. Mikami, K. Hiraki, S. Sakuma, T. Lino, F. Arai, T. Endo, Y. Fjiwaki, H. Fukuzawa, M. Hase, T. Hayakawa, K. Hiramatsu, Y. Hoshino, M. Inaba, T. Ito, H. Karakawa, Y. Kasai, K. Koizumi, S. W. Lee, C. Lei, M. Li, T. Maeno, S. Matsusaka, D. Murakami, A. Nakagawa, Y. Oguchi, M. Olikawa, T. Ota, K. Shiba, H. Shintaku, Y. Shirasaki, K. Suga, Y. Suzuki, N. Suzuki, Y. Tanaka, H. Tezuka, C. Toyokawa, Y. Yalikun, M. Yamada, M. Yamagishi, T. Yamano, A. Yasumoto, Y. Yatomi, M. Yazawa, D. D. Carlo, Y. Hosokawa, S. Uemura, Y. Ozeki, K. Goda., Intelligent Image-Activated Cell Sorting. *Cell*, 175, 1, 266-276 (2018).

- [46] B. Guo, C. Lei, H. Kobayashi, T. Ito, Y. Yalikun, Y. Jiang, Y. Tanaka, Y. Ozeki, K. Goda, High-throughput, label-free, single-cell, microalgal lipid screening by machine-learning-equipped optofluidic time-stretch quantitative phase microscopy. *Cytometry A*. 91, 494-502 (2017).
- [47] X. Huang, Y. Jiang, X. Liu, H. Xu, Z. Han, H. Rong, H. Yang, M. Yan, H. Yu, Machine Learning Based Single-Frame Super-Resolution Processing for Lensless Blood Cell Counting. *Sensors*. 16, 1836 (2016).
- [48] S. A. Damiani, D. Rossi, H. N. Joensson, S. Damiani, Artificial intelligence application for rapid fabrication of size-tunable PLGA microparticles in microfluidics. *Sci. Rep.*, 2020, 10, 19517 (2020).
- [49] Y. Mahdi, K. Daoud, "Microdroplet size prediction in microfluidic systems via artificial neural network modeling for water-in-oil emulsion formulation. *J. Dispers. Sci. Technol.* 38, 1501-1508 (2017).
- [50] D. Stoecklein, K. G. Lore, M. Davies, S. Sarkar, B. Ganapathysubramanian, Deep Learning for Flow Sculpting: Insights into Efficient Learning using Scientific Simulation Data. *Sci. Rep.*, 7, 46368 (2017).
- [51] Z. Zhou, X. Li, R. N. Zare, Optimizing Chemical Reactions with Deep Reinforcement Learning. *ACS Cent. Sci.*, 3, 1337-1344 (2017).
- [52] B. A. Rizkin, A. S. Shkolnik, N. J. Ferraro, R. L. Hartman, Combining automated microfluidic experimentation with machine learning for efficient polymerization design. *Nat. Mach. Intell.* 2020, 2, 200-209 (2020).
- [53] O. B. Mercan, V. Kılıç, M. Sen, Machine learning-based colorimetric determination of glucose in artificial saliva with different reagents using a smartphone coupled PAD. *Sens. Actuators B: Chem.* 329, 15, 129037 (2021).
- [54] R. S. Sutton, A. G. Barto, Reinforcement Learning: An Introduction. *MIT Press*, (1998).
- [55] X. Wang, S. Wang, X. Liang, D. Zhao, J. Huang, X. Xu, B. Dai, Q. Miao, Deep Reinforcement Learning: A Survey. *IEEE Trans Neural Netw Learn Syst. PP* (2022), doi:10.1109/TNNLS.2022.3207346.
- [56] S. Gronauer, K. Diepold, Multi-agent deep reinforcement learning: a survey. *Artificial Intelligence Review*. 55, 895–943 (2022).
- [57] J. Degraeve, F. Felici, J. Buchli, M. Neunert, B. Tracey, F. Carpanese, T. Ewalds, R. Hafner, A. Abdolmaleki, D. de Las Casas, C. Donner, L. Fritz, C. Galperti, A. Huber, J. Keeling, M. Tsimpoukelli, J. Kay, A. Merle, J.-M. Moret, S. Noury, F. Pesamosca, D. Pfau, O. Sauter, C. Sommariva, S. Coda, B. Duval, A. Fasoli, P. Kohli, K. Kavukcuoglu, D. Hassabis, M. Riedmiller, Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature*. 602, 414–419 (2022).
- [58] M. Cai, M. Hasanbeig, S. Xiao, A. Abate, Z. Kan, Modular Deep Reinforcement Learning for Continuous Motion Planning With Temporal Logic. *IEEE Robotics and Automation Letters*. 6, 7973–7980 (2021).

- [59] S. Backman, D. Lindmark, K. Bodin, M. Servin, J. Mörk, H. Löfgren, Continuous Control of an Underground Loader Using Deep Reinforcement Learning. *Machines*. 9, 216 (2021).
- [60] Y. Ji, J. Wang, J. Xu, D. Li, Data-Driven Online Energy Scheduling of a Microgrid Based on Deep Reinforcement Learning. *Energies*. 14, 2120 (2021).
- [61] Y. Wan, Z. Zhao, M. Yang, G. Xu, H. Ying, J. Wu, P. S. Yu, Improving automatic source code summarization via deep reinforcement learning. Proceedings of the 33rd ACM/IEEE International Conference on Automated Software Engineering (Association for Computing Machinery, New York, NY, USA, 2018; <https://doi.org/10.1145/3238147.3238206>), ASE 2018, pp. 397–407.
- [62] J. Luketina, N. Nardelli, G. Farquhar, J. Foerster, J. Andreas, E. Grefenstette, S. Whiteson, T. Rocktäschel, A Survey of Reinforcement Learning Informed by Natural Language. *arXiv [cs.LG]* (2019), (available at <http://arxiv.org/abs/1906.03926>).
- [63] K. Shao, Z. Tang, Y. Zhu, N. Li, D. Zhao, A Survey of Deep Reinforcement Learning in Video Games. *arXiv [cs.MA]* (2019), (available at <http://arxiv.org/abs/1912.10944>).
- [64] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, D. Hassabis, Human-level control through deep reinforcement learning. *Nature*, 518, 529-533 (2015).
- [65] P. R. Wurman, S. Barrett, K. Kawamoto, J. MacGlashan, K. Subramanian, T. J. Walsh, R. Capobianco, A. Devlic, F. Eckert, F. Fuchs, L. Gilpin, P. Khandelwal, V. Kompella, H. Lin, P. MacAlpine, D. Oller, T. Seno, C. Sherstan, M. D. Thomure, H. Aghabozorgi, L. Barrett, R. Douglas, D. Whitehead, P. Dürr, P. Stone, M. Spranger, H. Kitano, Outracing champion Gran Turismo drivers with deep reinforcement learning. *Nature*. 602, 223–228 (2022).
- [66] L. Kaiser, M. Babaeizadeh, P. Milos, B. Osinski, R. H. Campbell, K. Czechowski, D. Erhan, C. Finn, P. Kozakowski, S. Levine, A. Mohiuddin, R. Sepassi, G. Tucker, H. Michalewski, Model-Based Reinforcement Learning for Atari. *arXiv [cs.LG]* (2019), (available at <http://arxiv.org/abs/1903.00374>).
- [67] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, D. Hassabis, A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*. 362, 1140–1144 (2018).
- [68] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. V. D. Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature*. 529, 484-489 (2016).
- [69] Y. Zou, T. Chen, X. Chen, J. Li, Robotic seam tracking system combining convolution filter and deep reinforcement learning. *Mech. Syst. Signal Process.* 165, 108372 (2022).
- [70] X. Zhu, F. Zhang, H. Li, Swarm Deep Reinforcement Learning for Robotic Manipulation. *Procedia Comput. Sci.* 198, 472–479 (2022).

- [71] A. F. V. Muzio, M. R. O. A. Maximo, T. Yoneyama, Deep Reinforcement Learning for Humanoid Robot Behaviors. *J. Intell. Rob. Syst.* 105, 12 (2022).
- [72] J. Hua, L. Zeng, G. Li, Z. Ju, Learning for a Robot: Deep Reinforcement Learning, Imitation Learning, Transfer Learning. *Sensors* . 21 (2021), doi:10.3390/s21041278.
- [73] R. Liu, F. Nageotte, P. Zanne, M. de Mathelin, B. Dresch-Langley, Deep Reinforcement Learning for the Control of Robotic Manipulation: A Focussed Mini-Review. *Robotics*. 10, 22 (2021).
- [74] J. Ibarz, J. Tan, C. Finn, M. Kalakrishnan, P. Pastor, S. Levine, How to train your robot with deep reinforcement learning: lessons we have learned. *Int. J. Rob. Res.* 40, 698–721 (2021).
- [75] I. Akkaya, M. Andrychowicz, M. Chociej, M. Litwin, B. McGrew, A. Petron, A. Paino, M. Plappert, G. Powell, R. Ribas, Others, Solving rubik’s cube with a robot hand. *arXiv preprint arXiv:1910.07113* (2019) (available at <https://arxiv.org/abs/1910.07113>).
- [76] L. Pinto, M. Andrychowicz, P. Welinder, W. Zaremba, P. Abbeel, Asymmetric Actor Critic for Image-Based Robot Learning. *arXiv [cs.RO]* (2017), (available at <http://arxiv.org/abs/1710.06542>).
- [77] T.-C. Liang, Z. Zhong, Y. Bigdeli, T.-Y. Ho, K. Chakrabarty, R. Fair, Adaptive Droplet Routing in Digital Microfluidic Biochips Using Deep Reinforcement Learning. *PMLR*, 2020, 119, 6050-6060 (2020).
- [78] X. Y. Lee, A. Balu, D. Stoecklein, B. Ganapathysubramanian, S. Sarkar, A case study of deep reinforcement learning for engineering design: Application to microfluidic devices for flow sculpting. *J. Mech. Des.* 141, 11, 111401 (2019).
- [79] O. J. Dressler, P. D. Howes, J. Choo, A. J. deMello, Reinforcement Learning for Dynamic Microfluidic Control. *ACS Omega*. 3, 10084–10091 (2018).
- [80] A. T. H. Lau, H. M. Yip, K. C. C. Ng, X. Cui, R. H. W. Lam, Dynamics of Microvalve Operations in Integrated Microfluidics. *Micromachines*. 5, 1, 50-65 (2014).
- [81] J.-Y. Qian, C.-W. Hou, X.-J. Li, Z.-J. Jin, Actuation Mechanism of Microvalves: A Review. *Micromachines*. 11, 2, 172 (2020).
- [82] Y. Chen, Y. Tian, Z. Xu, X. Wang, S. Yu, L. Dong, Microfluidic droplet sorting using integrated bilayer micro-valves. *Appl. Phys. Lett.* 109, 143510 (2016).
- [83] T. V. Nguyen, P. N. Duncan, S. Ahrar, E. E. Hui, Semi-autonomous liquid handling via on-chip pneumatic digital logic. *Lab Chip*. 12, 3991–3994 (2012).
- [84] J. W. Hong, V. Studer, G. Hang, W. F. Anderson, S. R. Quake, A nanoliter-scale nucleic acid processor with parallel architecture. *Nature Biotechnology*. 22, 435-439 (2004).
- [85] M. A. Unger, H.-P. Chou, T. Thorsen, A. Scherer, S. R. Quake, Monolithic Microfabricated Valves and Pumps by Multilayer Soft Lithography. *Science*. 288, 113-116 (2000).
- [86] D. C. Duffy, J. C. McDonald, O. J. A. Schueller, G. M. Whitesides, Rapid Prototyping of Microfluidic Systems in Poly(dimethylsiloxane). *Anal. Chem.* 70, 4974-4984 (1998).
- [87] L.-S. Jang, Y.-J. Li, S.-J. Lin, Y.-C. Hsu, W.-S. Yao, M.-C. Tsai, C.-C. Hou, A stand-alone,

- peristaltic micropump based on piezoelectric actuation. *Biomed Microdevices*. 9, 185-194 (2007).
- [88] S. Okamoto, Y. Ukita, Reflow process using wax for fabricating curved shaped molds of PDMS microchannels and chambers. *Nanoscale Microscale Thermophys. Eng.* 8, 100055 (2020).
- [89] J.-Y. Qian, C.-W. Hou, X.-J. Li, Z.-J. Jin, Actuation Mechanism of Microvalves: A Review. *Micromachines*. 11 (2020).
- [90] Y. Li, C. Pan, Y. Li, E. Kumacheva, A. Ramachandran, An exploration of the reflow technique for the fabrication of an in vitro microvascular system to study occlusive clots. *Biomed. Microdevices*. 19, 82 (2017).
- [91] D. Kim, W. S. Lee, S. Shin, H. S. Rho, J. Dai, J. Y. Yun, J. W. Hong, Quantitative Analysis of Pneumatically Driven Biomimetic Micro Peristalsis. *Sci. Technol. Adv. Mater.* 6, 2428–2434 (2014).
- [92] Y.-C. Chen, G.-Y. Chen, Y.-C. Lin, G.-J. Wang, A lab-on-a-chip capillary network for red blood cell hydrodynamics. *Microfluid. Nanofluidics*. 9, 585–591 (2010).
- [93] P. M. Fordyce, C. A. Diaz-Botia, J. L. DeRisi, R. Gomez-Sjoberg, Systematic characterization of feature dimensions and closing pressures for microfluidic valves produced via photoresist reflow. *Lab Chip*. 12, 4287–4295 (2012).
- [94] T. Silva Santisteban, O. Rabajania, I. Kalinina, S. Robinson, M. Meier, Rapid spheroid clearing on a microfluidic chip. *Lab Chip*. 18, 153–161 (2017).
- [95] M. Kaminaga, T. Ishida, T. Omata, Fabrication of Pneumatic Microvalve for Tall Microchannel Using Inclined Lithography. *Micromachines*. 7 (2016).
- [96] S. Baig, G. Jiang, Q. Sun, M. R. Wang, Fabrication of single-mode channel polymer waveguides using vacuum assisted microfluidic soft lithography. *Journal of the European Optical Society - Rapid publications*. 8 (2013).
- [97] A. Flores, S. Song, S. Baig, M. R. Wang, Vacuum-Assisted Microfluidic Technique for Fabrication of Guided Wave Devices. *IEEE Photonics Technol. Lett.* 20, 1246–1248 (2008).
- [98] S.-J. Ahn, J. Moon, Vacuum-assisted microfluidic lithography of ceramic microstructures. *J. Am. Ceram. Soc.* 88, 1171–1174 (2005).
- [99] H. Yang, P. Deschatelets, S. T. Brittain, G. M. Whitesides, Fabrication of high performance ceramic microstructures from a polymeric precursor using soft lithography. *Adv. Mater.* 13, 54–58 (2001).
- [100] N. L. Jeon, I. S. Choi, B. Xu, G. M. Whitesides, Large-area patterning by vacuum-assisted micromolding. *Adv. Mater.* 11, 946–950 (1999).
- [101] H. Van Hasselt, A. Guez, D. Silver, Deep reinforcement learning with Double Q-learning. *Proc. Conf. AAAI Artif. Intell.* 30 (2016).
- [102] D. P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization. *arXiv [cs.LG]* (2014), (available at <http://arxiv.org/abs/1412.6980>).
- [103] M. Andrychowicz, F. Wolski, A. Ray, J. Schneider, R. Fong, P. Welinder, B. McGrew, J. Tobin,

P. Abbeel, W. Zaremba, Hindsight Experience Replay. *arXiv [cs.LG]* (2017), (available at <http://arxiv.org/abs/1707.01495>).

[104] K. Brower, A. K. White, P. M. Fordyce, Multi-step Variable Height Photolithography for Valved Multilayer Microfluidic Devices. *J. Vis. Exp.*, 119, 55276, (2017).

謝辞

本研究を遂行するにあたり、非常に熱心なご指導を賜りました浮田芳昭准教授に深く感謝いたします。また、本研究の討論の場で大変有益なご指導、ご助言を賜りました藤森篤教授、野田善之教授、石川陽教授、牧野浩二准教授、孕石泰丈准教授に深く感謝いたします。

そして、様々な議論を交わし、研究遂行に多大なご協力を頂いた浮田研究室の皆様から感謝の意を評します。