

深層強化学習を用いた集束超音波における焦点制御

日大生産工(院) ○佐藤 龍成 日大生産工 沖田 浩平

1. まえがき

近年, アレイトランスデューサをデバイスとして超音波を体内に集束させて治療する集束超音波 (HIFU: High intensity focused ultrasound) が注目され, メス等で身体に傷をつけることなく最小限かつ高効率な治療が可能となってきている. しかしながら, 人体の部位による超音波の屈折が発生し焦点を定めることが困難であることや, 焦点に発生したキャビテーション気泡によって音場が変化し, 集束超音波焦点がターゲットである病巣から外れる問題がある. これらの問題は, 複数の超音波振動素子からなるアレイトランスデューサの位相遅延量の制御によって解決可能であるが, 制御法は十分に確立されていない.

人工知能(AI)技術の進展によって様々な分野において広く応用されており, 医療業界にも利用されている. 特に機械学習(Machine Learning)は, データ整理や画像診断装置の精度向上に貢献しており, 既に機械学習を用いた診断支援機器が開発されている¹⁾. しかしながら, これまでの応用は主に診断支援や情報処理に留まっており, 機械学習を用いた制御に関する研究は行われていない. 機械学習が診断だけでなく治療における制御や最適化に関する研究は十分に行われていないのが現状である.

本研究では, フェーズドアレイトランスデューサの最適な制御方法への強化学習の適用に向けて, 音響不均質媒体中における強化学習による焦点制御の基礎検討として, 簡略化したモデルでその有効性について調べた.

2. 解析モデル

Fig.1にフェーズドアレイトランスデューサを用いた集束超音波の簡略化モデルを示す. このモデルでは, トランスデューサ素子に対応する音源より発生した超音波が一様な媒質中を伝搬して焦点に到達する状況を考える. 各素子から放出される音波の位相は, 超音波が焦点に集束するために調整されており, 音源から発する音波は次式で表される.

$$\sin(2\pi f(t + \tau)) \quad (1)$$

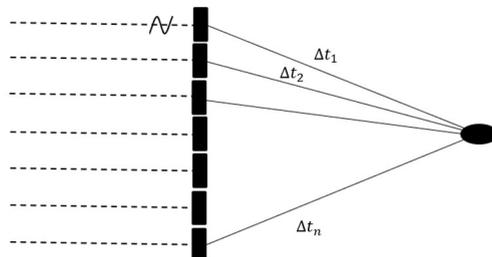


Fig.1 集束超音波の簡略化モデル

ここで, f は周波数, τ_n は各トランスデューサ素子に対して適応される位相遅延量を示している. 各音源から焦点までの伝搬時間を Δt_n とすると, 焦点における音圧は

$$\sum_{k=1}^n \sin(2\pi f(t + \tau_k - \Delta t_k)) \quad (2)$$

と表される. 焦点における振幅を最大化する位相遅延量 τ は次のようになる.

$$\tau_n = \frac{d}{c} - \frac{n}{f} \quad (3)$$

3. 深層強化学習

3.1 深層強化学習の概要と適用

強化学習は, 与えられた環境(状態)において学習者と呼ばれるエージェントが観測, 行動し, 目的とした価値(報酬)を最大化するまで試行錯誤し続ける学習方法である. Fig.2に強化学習の概念図を示す. 時点 t において, エージェントは環境から状態 S_t を観測し, 与えられた方策をもとに行動 A_t を出力する. 次に, エージェントの行動によって変動した環境は状態 S_{t+1} に更新され, それに応じた報酬 R_t をエージェントが得る. エージェントは報酬 R_t に基づいて方策の改善を繰り返し, 結果に基づいて方策を改善する²⁾.

本研究では, エージェントは行動 A_t として各超音波素子の位相遅延量 τ_n とし, 超音波の周波数 f に応じて $[0 \sim 1/f]$ の範囲で離散的に設定する. エージェントが観測する状態 S_t は, 各超

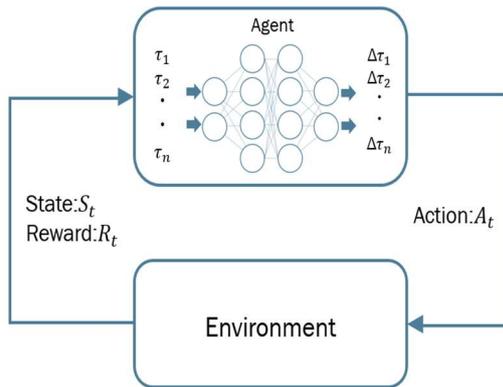


Fig.2 深層強化学習の概念図

音波素子に設定されている位相遅延量 τ である。報酬 R_t は行動 A_t に基づいて得られる振幅値の変化によって決定される。具体的には、焦点での振幅値が前回の振幅値と比較して増加した場合、報酬が与えられる。この構造により、エージェントは振幅値の最大化を目的とした最適な位相遅延量の調整を学習する。

3.2 DQN

DQNは、強化学習のQ学習に深層学習のニューラルネットワークを導入したアルゴリズムである。従来のQ学習では、状態と行動の組み合わせが増えると、膨大なQ値を管理する必要があり、対応できない場合がある。DQNは、ニューラルネットワークを用いて状態行動空間に対してQ値を近似し、効率的に最適な学習をするアルゴリズムである。DQNでは、強化学習のQ学習と同様に方策の更新式は次式で表される。

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha(r + \gamma * \max_{a'} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)) \quad (4)$$

右辺第1項は、状態 S_t における行動 A_t に対して見積もった価値である。第2項は、見積もった価値と実際の価値の誤差であり、損失関数表す。

4. 学習結果と考察

Fig.3に各エピソードの報酬を示す。学習が進むにつれて平均報酬が増加している。特に150エピソード目を境に上昇し始め、特定のエピソードで累積報酬が500に達していることから理想的な行動を選択したと見られる。しかし、報酬が収束せずに変動していることから学習が安定した段階には至っていないと考えられる。Fig.4に再学習における累積報酬の期待値を示す。累積報酬が最大の500に達していた

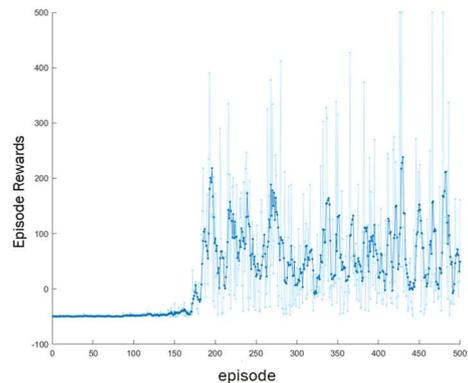


Fig.3 各エピソードの報酬

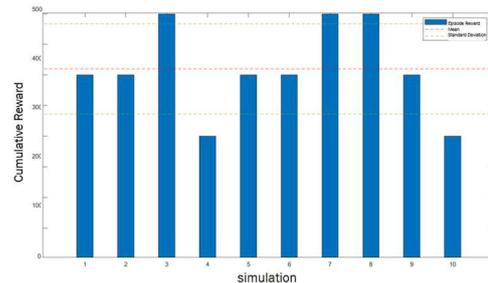


Fig.4 再学習における累積報酬の期待値

ものの安定した学習は見られなかった。報酬の定義が不十分であると考えられる。

5. 結言

本研究では、集束超音波治療におけるフェーズドアレイトランスデューサの最適な制御方法への深層強化学習の適応に向けて、簡略化モデルでその有効性について調べた。その結果、深層強化学習が位相遅延量の制御に対して一定の効果が得られることが示唆された。学習過程においては、期待される結果が安定して得られないケースがあり、学習精度の課題が明らかになった。

今後の課題は、環境設計における学習の安定性向上、超音波伝播モデルにMATLABのk-waveを用いて複雑な環境で適応することを目指す。

参考文献

- 1) T. Nakao, S. Hanaka, Y. Nomura, et al. : Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography, J. Magnetic Resonance Imaging, 47(4), 948-953,2018
- 2) 齋藤康毅, ゼロから作るDeep Learning4-強化学習編, 株式会社オライリー・ジャパン, (2022)