# AIによるAMD制御システム開発

青山優也

中 塚 光 -(設計本部)

### 鈴木貴博 (設計本部)

# Development of AMD Control System by AI

Yuya Aoyama

Koichi Nakatsuka

Takahiro Suzuki

# Abstract

In recent years, research and development to apply AI to the fields of structural design and building control has become more active. Reinforcement learning, which is a type of AI, is a framework for learning general control rules, including non-linear cases by trial and error, and has been successful in acquiring game strategies and controlling robots. In this paper, we report an attempt to learn the AMD control law for damping the walking vibration of an indoor bridge (crossing corridor) and wind vibration of a building by reinforcement learning. The simulation using the learned control law and the experiment applied to the actual system confirmed that the vibration control effect exceeded the control law by the conventional optimal control.

### 概 要

近年,構造設計や建築制御の分野にAIを適用しようとする研究開発が活発化している。AIの一種である強化 学習は非線形の場合を含む一般的な制御則を試行錯誤的に学習する枠組みであり,ゲーム戦略の獲得やロボッ トの制御で成功を収めている。本報においては,屋内ブリッジ(渡り廊下)の歩行振動および建物の風揺れを制 振対象としたAMDの制御則を,強化学習により獲得する試みについて報告する。学習した制御則を用いたシミュ レーション及び,実機システムに適用した実験により,従来の最適制御による制御則を上回る制振効果が得られ ることが明らかとなった。

### 1. はじめに

近年,AIの発展は目覚ましく,画像認識をはじめとし て産業界での実用例も増えている。建築構造分野におい ても,構造設計や建築制御に適用しようとする研究開発 が活発化している。AIの一種である強化学習<sup>11</sup>は非線形 の場合を含む一般的な制御則を試行錯誤的に学習する枠 組みであり,ゲーム戦略の獲得やロボットの制御で成功 を収めている。囲碁や将棋などのゲームにおいては人間 を上回る成績を出しているものも多く出てきているが, 発展途上の分野であり,建設業界のみならず産業界にお ける実用例はほとんどない。

一方,各種外乱に起因する建築物の水平振動や,床の 上下振動を低減させる有効な手法として,Active Mass Damper (AMD)が知られている。AMDは,対象構造物に 取り付けた重りを能動的に動かすことで対象構造物の振 動を低減する技術である。AMDの制御則として理論的知 見に基づいた多くの手法が提案されているが,中でも最 適制御と呼ばれる線形手法が広く実装されてきた<sup>2)</sup>。

大林組は,業界に先駆けて,AMDを対象とした制御則 を,強化学習により獲得する研究を進めてきた。強化学 習を用いることで、シミュレーションの段階で装置の特 性や実際の環境などを加味して学習することができる。 そのため、従来の理論に基づく方法を適用した場合より も、より実際の環境に適応した制御則の獲得を期待でき る。本報においては、ブリッジ(3.1節で詳述)の歩行振 動および建物の風揺れを対象として、シミュレーション におけるAMDの制御則の獲得および性能評価と、獲得し た制御則の実機への適用結果について報告する。

# 2. 強化学習の概要

### 2.1 強化学習

強化学習とは、エージェント(AI)が環境(制御対象) との相互作用を通じて最適な行動規則を獲得するための 枠組みである。強化学習の概要をFig. 1に示す。各時刻 t=0,1,2,…において、エージェントは環境の状態 $s_t$ を観測し、 方策 $\pi(a_t | s_t)$ (AIの行動規範)に従って行動 $a_t$ を選択す

エージェント (AI)		環境(制御対象)
<ul> <li>         ・方策π (AIの行動規範)         に従って行動a,を選択         </li> </ul>	行動a <sub>t</sub>	<ul> <li>・行動a<sub>t</sub>に対して状態s<sub>t</sub></li> <li>た。</li> </ul>
に従って行動加(こと)へ	「状態s <sub>t</sub>	之 S <sub>t+1</sub> (二 父 初)
・行動価値関数 $Q(s_t, a_t)$ を Bellman方程式により更新	報酬r <sub>t</sub>	・報酬r <sub>t</sub> を計算 
1	学習時のみ	



る。環境は現在の状態 $s_t$ と行動 $a_t$ に基づいて次の状態  $s_{t+1}$ に遷移し、エージェントは報酬 $r_t = r(s_t, a_t, s_{t+1})$ を受け取る。エージェントの目標は、割引率 $\gamma$ で割引いた将来に渡る累積報酬 $\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t, s_{t+1})$ の期待値を最大に する方策を獲得することである。

### 2.2 強化学習アルゴリズム

強化学習アルゴリズムは多数提案されているが、本報 では2.2.1~2.2.3項に示すQ学習、Deep Q-network (DQN), Actor-Critic法 (AC)を用いた3つの検討について報告する。 2.2.1 Q学習 状態 $s_t$ において行動 $a_t$ を選択したと きに将来に得られる累積報酬の期待値を、 $Q(s_t, a_t)$ で表 す。これは行動価値関数と呼ばれる。行動価値関数を環 境との相互作用から推定できれば、 $argmax_{a_t}Q(s_t, a_t)$ に よって最適な行動を選択できる。Q学習は行動価値関数 をテーブルで表現する。行動価値関数 $Q(s_t, a_t)$ はBellman 方程式に基づく(1)式の更新則によって推定される。

 $Q(s_t, a_t) = (1 - \alpha)Q(s_t, a_t) + \alpha \left\{ r_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) \right\} (1)$ ここで、  $\alpha (0 \le \alpha \le 1)$ は学習率である。

2.2.2 **Deep Q-network(DQN)** 状態が離散かつ有限 であれば行動価値関数をテーブルで表現できるが,状態 が連続の場合は関数近似器が必要である。Deep Qnetwork(DQN)では行動価値関数をニューラルネットで 近似し,ニューラルネットのパラメータθを(2)式で定義 する損失関数L(θ)の最小化によって推定する。

$$L(\theta) = E\left[\left\{r_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; \theta') - Q(s_t, a_t; \theta)\right\}^2\right] \quad (2)$$

ここで、 $\theta$ 'は更新前のパラメータであり、損失関数の 計算においては固定する。また、E[x]はxの期待値である。 2.2.3 Actor-Critic法(AC) DQNの行動選択は $a_t$ に関する行動価値関数の最大化を必要とするが、行動が 連続値を取る場合、この最適化計算は時間がかかる。こ の問題を避けるため、Actor-Critic法(AC)では行動価値関 数だけでなく方策  $\pi$  もニューラルネットで表現され、そ のパラメータは行動価値関数を最大化するように更新さ れる。行動価値関数の推定は、Q学習およびDQNの場合 と同様にBellman方程式に基づく方法によって行われる。 Actor-Criticは行動選択時の最大化計算が不要であり、連 続的な行動に適している。

# 3. ブリッジの振動制御シミュレーション

# 3.1 制御対象モデル

制御対象モデルは、大林組技術研究所本館テクノス テーションの連絡通路であるブリッジ<sup>3)</sup>(以降,ブリッ ジ)とする。ブリッジのスパンは14mであり、中央部にリ ニアモータで駆動するAMDが設置されている。ブリッジ および装置の概要をFig. 2に、諸元をTable 1にそれぞれ 示す。ブリッジは1次モードが支配的であるため、AMD のマスを含め2質点の振動モデルに置き換える。シミュ レーションは質点モデルの2階常微分方程式を解くこと により行った。

# 3.2 学習方法

2章に示すQ学習・DQN・ACを用いて制御則の学習を 行った。3.1節に示す制御対象モデルに対して,各種条件 をTable 2のように設定した。Q学習の制御力は-100N,0 N,100Nの3段階,DQNの制御力は-100N,-50N,0N,5 0N,100Nの5段階とした。ここで,DQNについては学習 を効率よく進めることを目的として,学習の進み具合に 応じて報酬r<sub>t</sub>を第1~第3世代の3段階に設定した。第1, 第2世代の学習エピソード数は10,000とし,50エピソード 毎に保存したモデルパラメータの中から累積報酬が最大 であったものを選択して,次世代の学習はそこから開始 した。一方で,第3世代の学習エピソード数は20,000とし, 累積報酬が最大であったものを採用した。なお,エピソー ドは1回の学習のことを,エピソード数は学習回数のこと をそれぞれ表している。

ACの制御力は,最大出力が100Nとなるように出力層 に非線形関数を適用した。学習エピソード数は50,000と し,1,000エピソード毎に保存したモデルパラメータの中 から累積報酬が最大であったものを採用した。

学習時の加振波は、インパルス波、実際の歩行時の計 測データから逆算した歩行加振波<sup>4)</sup>、ブリッジの固有周 波数と同じ周波数の正弦波の計3波とした。歩行加振波の 時刻歴波形をFig. 3に示す。







### 3.3 学習結果

ACのインパルス波に対する学習曲線をFig. 4に示す。 同図より,安定した学習を行えていることが確認できる。 一例として、インパルス波・DQNによる学習モデルの 50エピソード毎の評価値をFig. 5に示す。既に3.2節で述 べたように第1, 第2世代は累積報酬が最大であったもの を選択して、次世代の学習はそこから開始しており、第 2, 第3世代はそれぞれ6500, 10150エピソードから開始と なっている。同図中の上段は、居住性能評価指針5)の知 覚確率(黒線)および、それの20点移動平均(赤線)を 示す。下段は、ブリッジ変位の絶対値について、0秒以降、 0.12秒以降, 0.24秒以降の最大値を示す。なお、ブリッジ 固有周期の半周期が約0.12秒であり、0.12秒以降は2山目 以降の, 0.24秒以降は3山目以降の最大値を意図している。 同図より,学習が進むに従って,知覚確率や変位応答を 切り下げる良好な制御則を獲得できていることが分かる。

ラフにおいて、上段は2軸グラフであり、第1軸はブリッ ジ変位(mm)を, 第2軸はマス変位(mm)を示す。ブリッジ 変位に関しては、実線が制御した場合、点線がブリッジ のみ(非制振)の場合の応答をそれぞれ示す。また下段 は制御力(N)を示す。最初はマスがランダムに動きブ リッジを励振しているが、学習が進むにつれて、マスが ブリッジに対して遅れて動くことで応答低減するマスダ ンパーの挙動を獲得していることが確認できる。

#### 3.4 シミュレーションによる検証

-500 報酬

-1000 累積

-1500

強化学習により学習したモデル(強化学習モデル)の 制振性能を検証することを目的として、Table 3に示す8



つのモデルをシミュレーションにより比較した。強化学 習モデルの比較対象として,1.非制振,2.TMD,3.最適制 御(最適変位フィードバック制御則<sup>2</sup>))を用いた。ここ で,最適制御については,強化学習モデルと条件を揃え るために制御力を100Nで頭打ちにする処理を行った。各 種パラメータは事前試験により100Nを超えて不安定現 象が生じない程度とした。

3.4.1項においては、学習アルゴリズムの違いによる影響を確認するために、学習波をインパルス波に固定し、 学習アルゴリズムをQ学習・DQN・ACとしたモデルを比較した。一方、3.4.2項においては、学習波の違いによる 影響を確認するために、学習アルゴリズムをACに固定し、 学習波をインパルス波・歩行加振波・正弦波としたモデ ルを比較した。制振効果の評価には、1/3オクターブ分析 <sup>5)</sup>を用いた。

3.4.1 強化学習手法の比較 学習波をインパルス波 に固定し、学習アルゴリズムをQ学習・DQN・ACとした モデル(Table 3の4~6)および、比較用モデル(Table 3 の1~3)を比較した。インパルス波・歩行加振波入力時 の各種時刻歴応答波形をFig. 7に示す。ここで、最も特徴 をとらえた部分として、インパルス波応答は0~1秒、歩 行加振波応答は1~4秒を拡大して示している。また、ブ リッジ加速度の1/3オクターブ分析結果をFig. 8に示す。 ここで、同図中のV-〇〇は知覚確率を表している。

いずれの加振波に対しても、ブリッジの固有周波数で ある4Hz付近のピークは強化学習モデルが最適制御より も小さくなっており、最適制御を上回る制振性能が確認 できる。また、3種類の強化学習モデルの中ではDQNが最 もピークが小さい。他方で、Q学習・DQNでは制御力が 離散であることから、制御力の切り替わりに伴う加速度 波形の乱れが見られる。これにより、1/3オクターブ分析 の歩行加振波入力における高周波領域で増幅が見られる。 一方、最適制御・ACは制御力が連続であることからこの ような増幅は見られず、より安定した制御となっている。 3.4.2 学習波の影響 学習アルゴリズムを3.4.1項に おいて高周波領域で増幅が見られなかったACに固定し、 学習波をインパルス波・歩行加振波・正弦波としたモデ

モデル	学習アル ゴリズム	学習波	概要
1.非制振			マスを固定した場合
2.TMD	_		同質量のマスで,最適同調式による最 適剛性・減衰を与えたパッシブモデル
3.最適制 御			最適制御フィードバック制御則の変位 にかかる重み係数を調整したモデル
4.Q学習	Q学習		Q学習により学習した,7350ステップに おけるモデル
5.DQN	DQN	インパ ルス波	DQNにより学習した, 22000ステップに おけるモデル
6.AC(イ			AC法により学習した, 12000ステップ
ンパルス)			におけるモデル
7.AC(歩	AC	歩行加	AC法により学習した, 49000ステップ
行加振波)		振波	におけるモデル
8.AC(正		正弦波	AC法により学習した,41000ステップ
弦波)			におけるモデル

Table 3 比較モデル一覧 List of Comparison Models











ル (Table 3の6~8) および,比較用モデル (Table 3の1 ~3)を比較した。3.4.1項と同様に、インパルス波・歩行 加振波入力時の各種時刻歴応答波形をFig. 9に、ブリッ ジ加速度の1/3オクターブ分析結果をFig. 10にそれぞれ 示す。Fig. 10に示す1/3オクターブ分析より、正弦波で学 習したモデルが最もピークを抑えられており、前節のD ONよりも小さく抑えられていることが確認できる。

3.4.3 強化学習モデルの考察 Fig. 7およびFig. 9 に示す制御力より,強化学習モデルは最適制御よりも早 いタイミングで制御力を発揮していることが多く,結果 として振動を素早く抑えていることが確認できる。これ は,最適制御は線形かつ装置特性を考慮しない条件で制 御力を求めているのに対して,強化学習モデルは制御力 の非線形性や装置特性等の実際の環境を加味して学習し たため,これらに起因する各種制約の中で最大限制御効 果を発揮するように学習できたものと考えられる。

# 4. 実機による実験(ブリッジ)

### 4.1 実験の概要

3章で得られた強化学習モデルおよび比較用モデルを 実機に適用し、人力で加振することにより制振効果を確 認した。強化学習モデルは、Q学習・DQNおよび、ACの 中で最も制振性能が優れていた正弦波により学習したモ デルを用いた(Table 3の4・5・8)。実験システムの構成 をFig. 11に示す。実験は、以下の1)~5)をリアルタイム に0.01秒刻みで繰り返し実行することにより行った。

- 変位計によりマスとブリッジの相対変位を,加速度 計によりブリッジの絶対加速度を計測し、コント ローラに入力する。
- 2) 微積分等の演算により状態量(状態st)を計算する。
- 状態量を基に、コントローラが制御力(行動a<sub>t</sub>)を 計算する。
- 4) 計算した制御力をリニアモータに指令する。
- 5) リニアモータが制御力を発揮し、環境が変化する。

# 4.2 加力方法

比較のために,4.1節で述べた6つのモデルそれぞれに ついて,以下の1),2)による2ケースの加振を実施した。

- 1) 踵加振 ブリッジの中央部において, 踵で加振する。
   数秒おきに5回加振し,これを1セットとする。
- 歩行加振 ブリッジの端から端まで約14mを一人 で歩行する。1往復を1セットとする。

なお、TMDに関しては、最適剛性・最適減衰時の荷重 をリニアモータに指令することで模擬した。



Fig. 11 実験システムの構成 Experimental System Configuration











(When the Learning Wave is Changed)

試験においては、バラつきを低減するために、各ケース10セットずつ実施し、平均値を採用した。ここで、AC は別日に追加試験を行った都合上、その他の5つのモデル とは異なる人間による加振であり、5セットの平均である。

### 4.3 実験結果

計測したブリッジの加速度波形について,1/3オクター ブ分析を行った。ケース毎の平均値をプロットした結果 をFig.12に示す。

踵加振に対しては、Q学習・DQN・ACともに最適制御 を上回る結果であり、シミュレーションによる分析と同様の傾向であった。一方、歩行加振に対しては、DQN、 ACはシミュレーションによる分析と同様の傾向であっ たが、Q学習は最適制御よりも劣る結果であった。これ は、Q学習の制御力の切り替わりが多く、リニアモータの 制御力出力における時間遅れの影響や、積分時のノイズ 処理における時間遅れの影響を受けたものと考えられる。 また、Q学習・DQNは高周波の増幅が見られるのに対し て、ACは増幅が少ないことも、シミュレーションによる 分析と同様の傾向である。

### 5. シミュレーションによる学習(建物)

# 5.1 制御対象モデル

制御対象モデルは、27階建ての建物頂部にAMDを設置 したモデルを想定する。振動解析にあたり、Fig. 13に示 すとおり各層の質量・剛性を27質点、AMDのマスを1質 点、合計28質点の串団子モデルに置換する。また、学習 の安定性向上を目的として、建物部分を等価な1質点に縮 約し、AMDのマスを含め2質点の振動モデルに置換した モデルの学習も行う。1質点系に縮約した際の建物および AMDの諸元をTable4に示す。シミュレーションは、建物 部分の各質点に時刻歴の風圧力を作用させ2階常微分方 程式を数値解析で解くことにより行った。

### 5.2 学習方法

2.3節に示すACを用いて学習を行った。5.1節に示す制 御対象に対してTable 5のように各種条件を設定した。

ここで,装置の最大出力の制約を想定して,最大出力 100kNとなるように制御力の出力層に非線形関数を適用 した。学習ステップ数は100,000とし,1,000ステップ毎に 保存したモデルパラメータの中から累積報酬が最大で あったものを採用した。

学習に用いる時刻歴の風圧力は、3章において対象モデルの1次固有周期と一致する正弦波とすると最も制振性 能が優れていたことから、周期4秒の正弦波とした。

学習に用いた正弦波の時刻歴波形をFig. 14に示す。

また,学習曲線をFig. 15に示す。同図より,学習が安 定していることが確認できる。



# 6. シミュレーションによる検証(建物)

5.2節で得られた学習済みモデルの検証を行う。学習は 2質点・28質点モデルで行ったが、検証には28質点モデル を使用した。検証波は、学習に用いた正弦波に加えて、 本建物模型を用いた風洞実験により作成した風力時刻歴 波形を入力とする検証も行った。また、比較対象として 最適制御に検証波を入力するシミュレーションも併せて 行った。ここで、制御力は±100kNで頭打ちにすること で5.2節と同じ条件とした。建物頂部の最大加速度を横軸 に,建物頂部とマスの最大相対変位を縦軸にプロットし たグラフをFig. 16に示す。ここで、最適制御については 建物の変位にかかる重み係数qをq=10<sup>11.5</sup>~10<sup>15</sup>まで,指数 を0.5刻みで変化させた。同図中の緑色の点が0.5刻みの点 を、数値が指数を表す。同図より、2質点学習モデルの方 が28質点学習モデルよりも最大加速度を抑えられている ことが分かる。これは、2質点学習モデルの方が28質点学 習モデルよりも評価モデルが単純であり、学習の安定性 が高かったためと考えられる。さらに、2質点学習モデル は最適制御と同程度のマス変位で,qを変化させることで は実現できない加速度を実現できていることが分かる。



また,各種時刻歴波形のグラフをFig. 17に示す。時刻歴 波形については,学習済みモデルで性能の良い2質点学習 モデルと, Fig. 16でそれぞれの入力波において最も加速 度を抑えることができている最適制御のモデルの結果を 代表して示す。風洞実験波については,制御力が2質点学 習モデルの方が大きな制御力を出すことでマス変位も大 きくなり,頂部加速度のピークが抑えられた。

### 7. まとめ

ブリッジの歩行振動および建物の風揺れを制振対象と したAMDの制御則を,強化学習アルゴリズムであるQ学 習・DQN・ACにより学習した。また,学習した制御則の 有効性を,シミュレーションにより確認した。更に,ブ リッジについては実機に実装する実験を行い,制振性能 を確認した。得られた知見を以下に示す。

 Q学習・DQN・ACいずれにおいても学習が可能で あり、概ね最適制御の制振効果を上回る結果が得られることを確認した。これは、強化学習モデル が制御力の非線形性や装置特性等の実際の環境を 加味して学習したためであると考えられる。



Fig. 16 最大加速度-最大マス変位(左:正弦波入力 右:風洞実験波入力) Maximum Acceleration-Mass Displacement(Left: Sine Wave Input, Right: Wind Tunnel Experimental Wave Input)



Fig. 17 各種時刻歷応答波(左:正弦波入力 右:風洞実験波入力) Various Time History Response Waves (Left: Sine Wave Input, Right: Wind Tunnel Experimental Wave Input)

- 2) Q学習・DQNは制御力が離散であることに起因して高周波領域で増幅が見られるのに対して、ACは制御力が連続であることから、このような増幅は見られず、より安定した制御が可能であることを確認した。
- 3) 外乱が歩行振動や風圧力のようなランダムな波形であったとしても、対象モデルの1次固有周期と一致する正弦波を学習波としたモデルの制振性能が最適制御よりも優れていることを確認した。
- 4) 建物の場合において、対象モデルが28質点であっても、2質点モデルに縮約して学習した方が最大加速度を抑えることができ、同程度のマス変位でありながら最適制御では実現できない加速度を実現できることを確認した。

AIの一種である強化学習の研究および振動制御への 適用は黎明期にあり,既往制御の性能を大きく上回る全 く新しい制御を生み出す可能性を秘めている。一方で, AIについては思考プロセスがブラックボックスである 問題が指摘されており,高い安全性が求められる建築物 への恒久的な実適用にあたっては,詳細に安定性を検証 していく必要があると考える。今後も,AIをAMDの制御 に活用することに留まらず,構造設計分野や建築制御の 分野への継続的な研究と検証に取り組みたい。

# 謝辞

本研究の実施にあたり、Laboro.AIの椎橋徹夫CEOおよび、吉岡琢氏、中野達之氏より貴重なご意見・ご協力を 賜りました。ここに御礼申し上げます。

# 参考文献

- 1) 森村哲郎: 強化学習, 講談社, 305p., 2019.5
- 日本建築学会:建築構造物の振動制御入門,349p., 2010.12
- 3) 武田篤史,他:常温硬化型高じん性高強度モルタル 「スリムクリート®」の屋内ブリッジへの適用,大林 組技術研究所報, No.74, 10p., 2010.12
- 石川理都子,他:実測データに基づく歩行加振力評 価手法の提案,日本建築学会学術講演梗概集(関東), pp. 391-392,2011.9
- 5) 日本建築学会:建築物の振動に関する居住性能評価 指針・同解説, 132p., 2004.5