

可変長予測ホライズンを有するモデル予測制御を用いた車両制御

M2021SC001 伏屋主水之介

指導教員：坂本登

1 はじめに

近年、様々な工業分野で自動化が推奨されており、数多くある自動化技術の中でも、モデル予測制御は汎用性の高さから幅広い分野で応用が期待される。応用先については文献 [1] で、石油精製、化学産業、発電プラント、鉄鋼プロセス、セメントプロセス、ごみ焼却炉、上下水道、空調制御、自動車産業、医療分野などが記述されており、益々多くの活用が見込まれる。モデル予測制御は有限な評価区間（ホライズン）において制御対象の未来の状態を予測し、評価区間内で目的関数を最小にする入力を求める制御手法である。

モデル予測制御は計算負荷が多いという短所がある。特に車両の経路生成に活用する場合には、障害物の移動も考慮する必要があり、その際に予測ホライズンを延ばすほど不確かな計算に時間を割くことになる。

計算時間の課題を解決する手法として文献 [2, 3] で利用されている C/GMRES 法が挙げられる。C/GMRES 法とは連続法 (Continuation method) と一般化最小残差法 (Generalized Minimal Residual Method) を組み合わせた手法である。この手法を用いることにより高速で最適制御入力列を解くことができる。だが文献 [4] で指摘されているように C/GMRES 法は連続変形法を利用しているため、制御車両の前方に他車両が割り込むことなどにより、最適解が急激に変化する場合には解を導くことができない可能性がある。

そこで予測ホライズンを工夫することで計算量を削減しながら適切な制御を行うことを考える。文献 [5] では可変長予測ホライズンを有するモデル予測制御において、ロバスト性を確保しながら計算量を削減するアルゴリズムに関して記述している。文献 [6] では可変長予測ホライズンのホライズン長を粒子群最適化法を用いることでホライズン長の最適化を行っている。文献 [7] では可変長予測ホライズンを人物追跡に使用しており、過去に人物が進行方向を変更した確率をオンラインで算出して予測モデルの正確さを判断し、これを可変長予測ホライズンに利用している。

本研究では二輪車両ロボットの経路生成において上記のように可変長予測ホライズンを利用することで計算時間の課題解決を試みる。周囲環境や車両の状態から強化学習により予測ホライズンの長さを学習させることで各時刻の状況に合わせたホライズン長に可変させる。

また本研究の最終目標は、人と空間を共有する場所で稼働する車両の制御であるため、移動する障害物が人である場合にも適切に経路を生成する必要がある。そのため人の周囲にパーソナルスペースを生成し、回避する様な経路で車両を移動させることを目標とする。

最後に、今回は文献 [8, 9, 10, 11] を参考に MATLAB でシミュレーション環境を構築している。

2 問題設定

本研究では港や倉庫、屋内での運搬車両など人との共存が必要となる車両の制御を考える。スタートからゴールまで可変長の予測ホライズンを採用したモデル予測制御を使用し適切な経路を生成する。

シミュレーション環境のイメージを図 1 に示す。これは車両と周囲環境を真上から見たものである。車両が移動する環境は障害物のある環境であり、障害物には移動しない障害物と人に見たてた移動する障害物の二種類が存在する。移動する障害物（人）にはパーソナルスペースを付与することで適切な経路生成を行うようにする。また、車両から出ている黒色の矢印が車両の前方を示している。

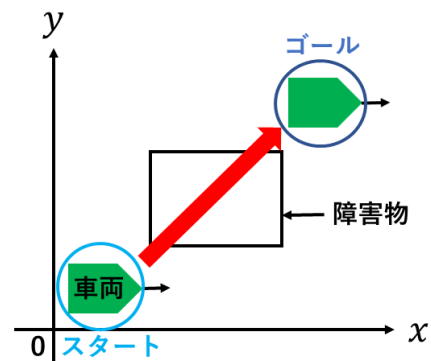


図 1 シミュレーション環境

3 車両モデル

本研究では可変長予測ホライズンを有するモデル予測制御をシミュレーション上で行う。以下の図 2 にシミュレーションで使用する車両モデルを示す。 u_1, u_2 は左右の車輪への入力である回転角速度、 r は車輪の半径、 θ は車両の旋回角度、 d は車幅である。左右の車輪の速度は

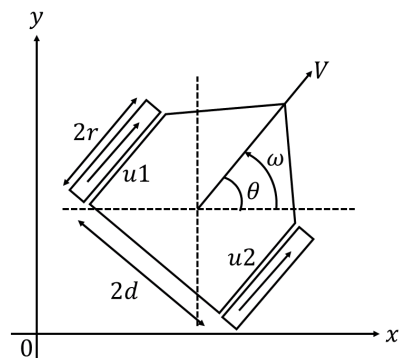


図 2 車両モデル

それぞれ以下の式 V_l, V_r で与えられる：

$$V_l = ru_1, \quad (1)$$

$$V_r = ru_2. \quad (2)$$

車両の重心速度 V と旋回角速度 ω は以下の式で与えられる：

$$V = \frac{(V_r - V_l)}{2}, \quad (3)$$

$$\omega = \frac{V_r - V_l}{2d}. \quad (4)$$

式 (1), (2) を式 (3), (4) に代入すると

$$V = \frac{r(u_2 - u_1)}{2}, \quad (5)$$

$$\omega = \frac{r(u_2 - u_1)}{2d}. \quad (6)$$

が得られる。式 (5) は車体の速度と左右の車輪の回転角速度（入力）との関係を表している。式 (6) は車体の旋回角速度と左右の車輪の回転角速度（入力）との関係を表している。

また、車両を真上から見た平面上に xy 直交座標をとり状態空間表現を考えると二輪車両ロボットは以下の式で表現できる：

$$\frac{d}{dt} \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} = \begin{bmatrix} \cos \theta \\ \sin \theta \\ 0 \end{bmatrix} V + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \omega.$$

4 目的関数

モデル予測制御では目的関数を使用して車両の状態や周囲環境を評価する。今回は目的関数が最小となる左右の車輪の回転角速度（入力）を求めることで二輪車両ロボットを制御する。また、本研究では車両の移動に人工ポテンシャル法を取り入れることでモデル予測制御を行っている。

目的関数は車両の現在位置とゴールの距離に関する項 J_1 、車両の現在位置と障害物までの距離の項 J_2 、左右の車輪の回転角速度（入力）に関する項 J_3 、車両の現在角度とゴールでの目標角度の差に関する項 J_4 からなる。ステップは刻み幅 h でシミュレーション区間を分割したものであり、現在時刻までのステップ数を n 、予測ステップ数を p とする。

J_1 の式を示す。 x_g, y_g はゴール座標、 x, y は車両の現在座標を示している。 $x(k|t)$ は時刻 t で予測した k ステップ目における車両の x 座標であり、 $y(k|t)$ は時刻 t で予測した k ステップ目における車両の y 座標である。 $x(k|t), y(k|t)$ をそれぞれ x, y で置く。また、 C_1 は重み係数であり正の定数である。

$$J_1 = \sum_{k=t}^{t+p-1} C_1 \sqrt{(x - x_g)^2 + (y - y_g)^2}.$$

J_2 の式を示す。 x_{ob}, y_{ob} は障害物の座標であり、 i は検知した個数である。 C_2 は重み係数であり正の定数である。

$$J_2 = \sum_{i=0}^{l-1} \sum_{k=t}^{t+p-1} C_2 \exp(-\sqrt{(x - x_{obs_i})^2 + (y - y_{obs_i})^2}).$$

J_3 の式を示す。 u は左右の車輪の回転角速度（入力） u_1, u_2 から成るベクトルであり、 $u(k|t)$ は時刻 t で予測した k ステップ目における入力である。 R は重みの行列である。

$$u(k|t) = \begin{bmatrix} u_1(k|t) \\ u_2(k|t) \end{bmatrix},$$

$$J_3 = \sum_{k=t}^{t+p-1} u(k|t)^T R u(k|t).$$

J_4 の式を示す。 $\theta(k|t)$ は、時刻 t で予測した k ステップ目における車両の角度であり、角度は x 軸正方向を基準に測るものとする。 $\theta_0(k|t)$ は時刻 t で予測した k ステップ目における目標地点へ向かう方向を表している。また、 $\theta(k|t), \theta_0(k|t)$ をそれぞれ θ, θ_0 で置く。 C_3, L_1 は重み係数であり、正の定数である。

$$J_4 = \sum_{k=t}^{t+p-1} (\theta - \theta_0)^2 C_3 \exp(-\frac{\sqrt{(x - x_g)^2 + (y - y_g)^2}}{L_1}).$$

5 パーソナルスペース

人が互いに向き合い、歩行している時に衝突しないよう方向転換をする距離がある。これはそれぞれ個人のパーソナルスペースに侵入し、回避行動をとり始める距離である。このように人は衝突する前にパーソナルスペースを考慮して回避行動を行っている。本研究では二輪車両ロボットの障害物回避において、人に見たてた移動する障害物の周囲にパーソナルスペースを付与することで、人に不快感を与えずより安全に回避行動を行うことを考える。また、パーソナルスペースは文献 [12] を参考に作成し、作成したパーソナルスペースを図 3 に示す。パーソナルスペースの範囲は点 A, B, C, D, E, F で構成される楕円内部であり、人の範囲は点 G を中心して描かれている円の内部である。 $\angle AFE$ も $\angle AFB$ と同じ角度 α である。

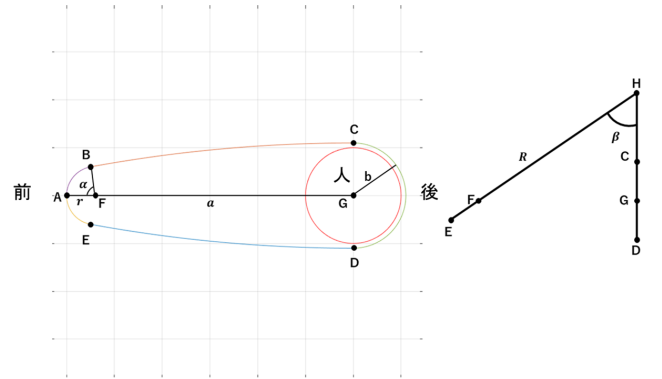


図 3 パーソナルスペース

以下にパーソナルスペースを構築する際に必要なそれぞれの点と角度の関係式を示す。

$$\begin{aligned}\overline{GA} &= a, \overline{GC} = b, \\ r &= \frac{\sqrt{a^2 + b^2}(\sqrt{a^2 + b^2} - (a - b))}{2a}, \\ \alpha &= \tan^{-1}\left(\frac{a}{b}\right), \beta = \tan^{-1}\left(\frac{b}{a}\right), \\ R &= \frac{\sqrt{a^2 + b^2}(\sqrt{a^2 + b^2} + a - b)}{2b}, \\ \overline{GH} &= R - b.\end{aligned}$$

上記の式より構築したパーソナルスペースと車両からの距離が一番近い点 (x_{ps}, y_{ps}) との距離を車両とパーソナルスペースの距離とする。この定義により、目的関数の車両とパーソナルスペースの距離の項は以下の式となる。また、 C_4 は重み係数であり、正の定数である：

$$J_5 = \sum_{i=0}^{l-1} \sum_{k=t}^{t+p-1} C_4 \exp(-\sqrt{(x - x_{ps_i})^2 + (y - y_{ps_i})^2}).$$

6 モデル予測制御の検証

二輪車両ロボットをモデル予測制御でシミュレーションした結果を図4,5に示す。このシミュレーションでは、スタートからゴールまで障害物を回避しながら二輪車両ロボットが移動する。図4はパーソナルスペースを考慮しない目的関数 $J_1 + J_2 + J_3 + J_4$ でシミュレーションをした結果であり、図5はパーソナルスペースを考慮する目的関数 $J_1 + J_2 + J_3 + J_4 + J_5$ でシミュレーションをした結果である。左から右に移動している箱状の物体が二輪車両ロボットである。緑色の円が連なっているのは人の移動した推移であり、黒色の円は移動しない障害物である。

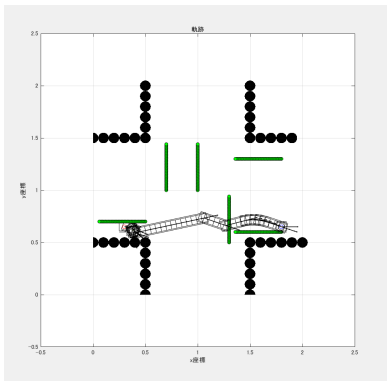


図4 パーソナルスペースなし

図4では二輪車両ロボットが障害物に当たる直前まで接近してから進路変更していることがわかる。図5では二輪車両ロボットは障害物に接近することなくゴールに到達していることがわかる。二つの経路の比較からパーソナルスペースを考慮する効果が見られる。

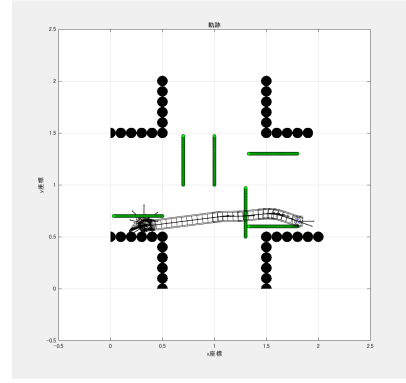


図5 パーソナルスペースあり

7 強化学習による可変長予測ホライズンの学習

強化学習により予測ホライズン長の学習を行う。モデル予測制御は離散時間における制御手法であるため、DQN (Deep Q-Network) を使用して学習を行うが、学習初期段階であるため確認として報酬はホライズンの長さのみとなっている。

本研究で使用するDQNのアルゴリズムについて説明する。DQNはQ学習にDNN (Deep Neural Network) を組み合わせたものであり、Q学習で用いられているテーブル形式の学習よりも効率的に学習が可能である。以下にDQNで使用されるアルゴリズムを記述する。

$$Q^\pi(s_t, a_t) \leftarrow Q^\pi(s_t, a_t) + \alpha(Y_t - Q^\pi(s_t, a_t)), \quad (7)$$

$$Y_t = r_{t+1} + \gamma \max_a Q(s_{t+1}, a'; \theta'). \quad (8)$$

Q は予測した期待値であり α は学習率である。 s_t, s_{t+1} はそれぞれの時刻 $t, t+1$ における状態を表しており、 a_t は時刻 t におけるアクションを表している。

また $Y_t - Q^\pi(s_t, a_t)$ はTD誤差といい、実際の期待値 (TDターゲット: Y_t) と予測した期待値の差をとるものである。 r_{t+1} は時刻 $t+1$ における報酬であり、 γ は割引率である。

$\max_a Q(s_{t+1}, a'; \theta')$ では、状態 s_{t+1} においてアクション a' を選択した場合に最大の期待値となるようなものを選択する。 θ' は学習パラメータである。またパラメータ更新は以下の損失関数で行う。

$$L(\theta) = E[(Y_t - Q(s_t, a_t; \theta))^2]. \quad (9)$$

式(8),(9)を利用して式(7)を更新することで期待値を更新していく。この期待値が大きくなるほど良いため、DQNでは期待値であるQ値を最大にするように学習していく。

図6,7に学習結果を示す。オレンジ色のグラフは人の検知に関するグラフであり、人を検出していない場合には1、検出した場合には0を示す。青色のグラフは可変長予測ホライズンの時刻ごとの長さであり、3から6までの長さで変化するように設定してある。また今回のシミュレーションでは、二輪車両ロボットのスタートとゴール

の位置をランダムに設定しているため、シミュレーション毎にシミュレーション終了時間が異なっている。

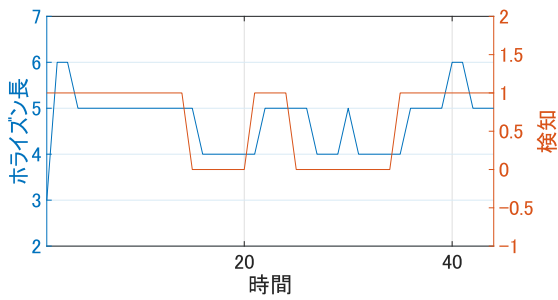


図 6 DQN での学習結果

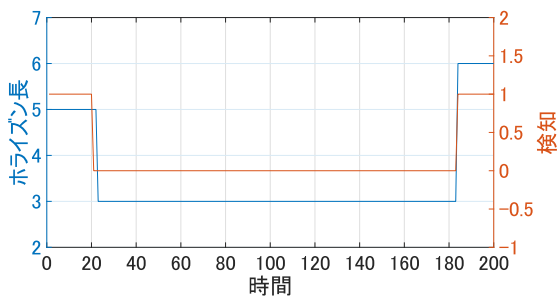


図 7 DQN での学習結果 2

図 6 は、ホライズン長に応じた報酬 p_t だけを与えた。以下に報酬関数の式を示す。

$$r_t = p_t. \quad (10)$$

人を検知した場合にはホライズン長は短めの 3, 4 になっており、人を検知していない場合にはホライズン長は長めの 5, 6 となっている。これは人を検知した場合にはホライズン長を短くすることで計算負荷を軽減し即応性を高め、人を検知していない場合にはホライズン長を長くして正確な制御を行っている。

図 7 では先ほどの報酬に加えて、時間ごとでの進んだ距離 dis_t 、障害物との距離 obs_t 、シミュレーション終了時間 ts_t を追加した。以下に報酬関数の式を示す。

$$r_t = p_t + dis_t - obs_t + ts_t. \quad (11)$$

結果としてシミュレーション開始時の不確定要素の多い場面ではホライズン長を二番目の長さである 5 を選択しており、ゴールが近づく終盤では不確定要素の発生区間が短いためホライズン長を最長の 6 にしていることがわかり、適切なホライズンの切り替えが行われている。ただし障害物検知時のホライズン長の可変はできていないため今後の課題となる。

8 おわりに

本研究では可変長予測ホライズンを有するモデル予測制御を使用し、人のパーソナルスペースを考慮した二輪

車両ロボットの経路生成を行った。パーソナルスペースを生成し、モデル予測制御で使用する目的関数を作成することで経路生成におけるパーソナルスペースの効果を確認した。また、DQN で強化学習を行うために簡潔な報酬で学習を実行した。その結果、学習で障害物を検知する場合のホライズンを短くし、障害物を検知していない場合にホライズンを長くするという方向性を確認することができた。しかし図 6, 7 は狙い通りの結果が出た場合であり、実行回数に対するホライズン可変の精度が低いため、今後は精度向上を考えて行う必要がある。

参考文献

- [1] 新見和久：『モデル予測制御の産業応用への展望』。計測自動制御学会 計測と制御 43 巻, 2004.
- [2] Toshiyuki Ohtsuka : 『A continuation/GMRES method for fast computation of nonlinear receding horizon control』, Automatica 40, 563/574 (2004).
- [3] 大山博之・山北昌毅：『連続法へ数値最適化手法を応用した並列非線形モデル予測制御』。第 57 回自動制御連合講演会, 群馬, 2014.
- [4] 川邊武俊：『道路情報と地形情報を利用した電気自動車の航続距離延長制御の研究』。科学研究費助成事業研究成果報告書, 平成 30 年 6 月 21 日
- [5] Rohan. C. Shekhar : 『Variable Horizon Model Predictive Control: Robustness Optimality』。Pembroke College. 2012.
- [6] Huiran Wang・Qidong Wang・Wuwei Chen・Linfeng Zhao・Dongkui Tan : 『Path tracking based on model predictive control with variable predictive horizon』。SAGE. 2021.
- [7] 武村紀子・中村泰・松本吉央・石黒浩：『可変長予測に基づく人物追跡』。情報処理学会関西支部大会。平成 21 年度。
- [8] ヤン. M. マチエヨフスキー (著)・Jan M. Maciejowski (原著)・足立修一 (翻訳)・菅野政明 (翻訳) 『モデル予測制御：制約のもとでの最適制御』。東京電機大学出版局。2005.
- [9] 真野翼・伏屋主水之介・大石泰章：『モデル予測制御を用いた二輪車両ロボットの障害物回避』。ロボティクス・メカトロニクス 講演会。2021.
- [10] 足立修一：『モデル予測制御の基礎』。日本ロボット学会誌。2014.
- [11] 小山健太郎・野中謙一郎：『障害物回避と切り返し点の自動調整によるモデル予測車庫入れ制御』。計測自動制御学会論文集。2014.
- [12] 劉建宏・大枝良直・角知憲：『パーソナルスペースを用いた障害物を回避する歩行者の群集流動』。土木学会論文集 D。2008.