

仮想工場における多数AGVの自律運転に関する研究

旭川高専 ○池田将晴 古川正志 渡辺美知子

要旨

自律分散型の生産システムを確立する上で、AGVの自律運転は重要な役割を持っている。本研究では、多数AGVが工場内で学習に基づき、自律運転を可能とすることを目的とした経路獲得と衝突回避の方法を報告する。

1. はじめに

近年、自律分散型生産システムの構築において、各ショップ間を自由に移動できるAGVの自律行動をどのように獲得できるかは大きな問題であり、多数の研究が行われている。この問題は、1)目的地までの経路獲得問題と、2)他のAGVとの協調行動問題の2つからなる。本研究ではコンカレントにこれらを実現するために1)はAGVの局所的な風景を簡易SDMを用いて獲得し、2)と並行してQ-学習を行い、行動を獲得する方法を提案し、その結果を報告する。

2. 簡易SDMによるシーン獲得

本研究では、経路獲得問題をシーン獲得による経路決定法として取り扱い、その学習による解法を提案する。

2.1 前提条件

AGVには以下の条件を与える。

1. AGVは周辺7×7マスの障害物の有無を検出できる。
2. AGVは通過点のシーンを記憶することができる。
3. AGVは現在地から目的（作業機械）の距離を測定できる。

2.2 問題の記述

先に設定した前提条件の下、問題を次のように設定する。

$$\begin{aligned} \min_{Seq(S_c), n} & W_1 F_1(Seq(S_c)) + W_2 F_2(n) \\ \text{subject to} & \\ & F_1(Seq(S_c)) \leq Limit \end{aligned} \quad (1)$$

ここで、 $S_c \in Scene$ はAGVが獲得したシーン、 $Scene$ はシーンの全集合、 $Seq(S_c)$ は一連のシーン、 $F_1(Seq(S_c))$ は経路の最大滞留時間、 $F_2(n)$ はシーンの数 n の評価、 $Limit$ は許容学習時間である。また、 w_1, w_2 は評価のウェイトである。

3. Q-学習の適応による経路獲得

3.1 SDMの簡略化

SDM(疎分布記憶)は一般化、抽象化、連想を実現するため、記憶データを2進ベクトルとして格納する。この2進ベクトルは「記憶装置内のアドレス」、または「そのアドレスか別のアドレスに格納されたデータ」の両方が可能なアドレス表現を用いる。これにより、人間の思い起こす能力をハード上にカウンターにより実現し、多大な記憶装置の場所を必要としない疎分布な記憶形態を指定したハミング距離で実現する。

3.2 シーン獲得法

AGVは検出した風景を2進化したベクトルとして記憶する。そして検出直前と直後のシーンベクトルのハミング距離を計算し、ある既定値をこえたとき、そのシーンを新たなシーンとして記憶する。この手続きを目的地到達まで続ける。

3.3 経路獲得アルゴリズム

Q-学習における状態 s を簡易SDMによって獲得されたシーン、行動 a を右旋回、左旋回、直進の3種の方向転換行動とし、この s - a 間にQ値(利得の期待値)を設定してQ-学習を適応する。Q値からの行動決定関数にはボルツマン分布

を採用し、確率的に決定する。

また、環境からの報酬は現在地からゴールまでの距離で評価する。さらに、経路は一連のシーンのつながりであることを考慮し、バケットブリゲードによりQ-学習の報酬 f_c を割引係数 Γ 、割引率 d_c を用いて以下のように与える

$$\Gamma_{j+1} = \Gamma_j \cdot d_c \quad (2)$$

$$f_{cj}(\text{直接評価}) = R(\text{成功}) \cdot \Gamma_j$$

$$f_{cj}(\text{直接評価}) = P(\text{失敗}) \cdot \Gamma_j \quad \text{但し, } j=1,2,\dots,RMS$$

ここでRMSは遡る過去のシーン数である。また、Q-学習の更新式は以下を採用する。

$$Q_{i+1}(s_t, a_t) = (1 - \alpha)Q_i(s_t, a_t) + \alpha \left[f_c + \gamma \max_{b \in A} Q_i(s_{t+1}, b) \right] \quad (3)$$

ここで、 α は学習率、 γ は減衰率、 A は行動集合である。

3.4 シミュレーション

実際のシミュレーションによる経路獲得の学習の様子を図2に示す。

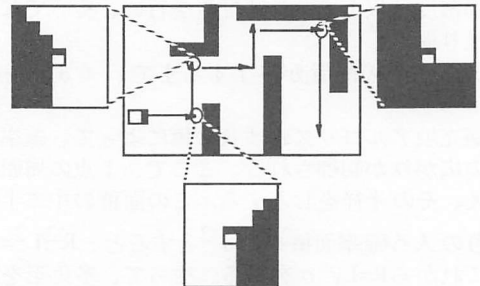


Fig.1 シーン獲得

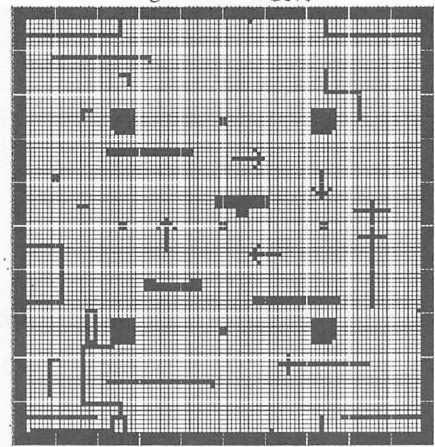


Fig.2 シミュレーション

(左下格納庫から作業機械までの経路)

4. 通信プロトコルの獲得

4.1 前提条件

シーン獲得について、通信プロトコルには次の条件を設定する。

1. AGVは次行程の作業装置の方角を知覚できる。
2. AGVは自分の進行方向に隣接する他のAGVについてのみ、その進行方向を知覚できる。
3. 作業機械は、その加工時間を0、AGVの収納数とする。

4.2 問題の記述

このような前提条件の下、問題を次のように設定する。

$$\begin{aligned} & \min_{Seq(A_i)} \max_i F_i(Seq(A_i)) \\ & \text{subject to} \\ & Seq(A_i) \supset M(A_i) \end{aligned} \quad (4)$$

ただし、

- A_i : i 番目のAGV
- $Seq(A_i)$: A_i の一連の行動
- $F(Seq(A_i))$: A_i の最大滞留時間
- $M(A_i)$: A_i の立ち寄る作業機械のシーケンス

とする。

5. オートマトンによる学習モデル

AGV 1台を1個のオートマトン A_i とみなすと、全AGVの集合は集団オートマトン A として、

$$A = \{A_i; i=1,2,\dots,TT\} \quad (5)$$

と記述する。ここで、 TT はAGVの総数である。また、個々のオートマトン A_i は、

$$A_i = \{I_i, O_i, S_i, F_i, G_i\} \quad (6)$$

の5項組で記述できる。個々で、各記号は、

- I_i : 入力 O_i : 出力
- S_i : AGVの状態
- F_i : 状態変位関数
- G_i : 出力関数

である。

5.1 入力 I_i

入力は以下のように構成される。

$$I_i(t) = \{P_i(t), M_i(t), P_j(t), D_j(t), E_j(t)\} \quad (7)$$

ここで、 t は時刻を表し、

- $P_i(t)$: AGVの現在位置(x,y)
- $M_i(t)$: 目標とする作業機械の位置(x,y)
- $P_j(t)$: 遭遇したAGVの座標値(x,y)
- $D_j(t)$: 遭遇したAGVの進行方向
- $E(t)$: AGVと遭遇したかの可否を表す変数

である。

AGVの進行方向は

$$D_i(t) = \{D \mid \text{North, East, South, West} : \text{北, 東, 南, 西}\} \quad (8)$$

とする。

遭遇可否変数 $E(t)$ は他のAGVと遭遇したとき $E(t)=1$ とし、遭遇していないときは $E(t)=0$ とする。

5.2 出力 O_i

出力はAGVの行動とする。時刻 t におけるこの行動を $O_i(t)$ とすると

$$O_i(t) = \{O_i \mid \text{Forward, Right, Left}\} \quad (9)$$

AGVは3方向、前進、右、左へ移動することができる。

これにより時刻 t におけるフィールド空間の進行座標方向 $D_j(t) = \{D_j \mid \text{North, East, South, West} : \text{北, 東, 南, 西}\} \quad (10)$

の4方向が決定される。

5.3 状態 S_i

$E(t)=1$ の時のAGVの状態は以下のように表される。

$$S_i = \{\text{Same, Different}\} \quad (11)$$

AGV、 A_i の状態がSameであるとき、遭遇したAGV、 A_i

との進行方向の状態が同方向 ($D_i(t)=D_j(t)$)であることを表す。それに対して状態がDifferentであるときは、遭遇したAGV、 A_i との進行方向の状態が異方向 ($D_i(t) \neq D_j(t)$)であることを表す。

5.4 状態遷移

AGVの状態は他のAGVと遭遇するたびに遷移する。ここではQ-学習を用い、状態遷移関数 F_i は式(3)を採用する。

5.5 出力関数 g_i

出力関数 g_i はAGVの現在の状態 S からその時点での行動 O を確率的に定める関数である。本アルゴリズムでは、行動決定関数はボルツマン分布により確率的に決定する。

6. Q-学習による通信プロトコルの獲得法

通信プロトコルとは複数のAGVがセンサ内に感知される相互の行動を理解し、共有の目的を果たす行動知識をいう。

6.1 Q-学習の設計

今、2台のAGV (A_i, A_j)が隣接したとする。時刻 t におけるこれらの行動をそれぞれ O_i, O_j とすればそれらの進行方向 $D_i(t), D_j(t)$ が定まる。これらを移動ベクトルとみなし、それぞれのAGVの位置ベクトルは

$$P_i(t+1) = P_i(t) + D_i(t), \quad P_j(t+1) = P_j(t) + D_j(t) \quad (12)$$

と予測できる。

この位置計算を用いて、相手の速度の状態が同方向 ($D_i(t)=D_j(t)$)、異方向 ($D_i(t) \neq D_j(t)$) と行動 O

(Forward, Right, Left) のそれぞれの状態の場合にこのQ値を設定する。Q-学習に必要な直接利益を以下のように設定する。

$$\begin{aligned} & \text{if } [P_i(t+1) \neq P_j(t+1) \cap P_i(t+1) \neq \nabla P_k(t)] \\ & \quad \text{then } C_i = \text{Reward and } C_j = \text{Reward} \\ & \text{elseif } [P_i(t+1) = P_j(t+1) \cup P_i(t+1) = \nabla P_k(t)] \\ & \quad \text{then } C_i = \text{Penalty and } C_j = \text{Penalty} \\ & \text{else} \\ & \quad C_i = 0 \end{aligned} \quad (13)$$

この計算では最初の判定が衝突回避成功、2番目が失敗を表している。

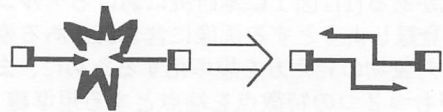


Fig.3 衝突回避行動の一例

7. おわりに

本報告では、簡易SDMを用いたシーン獲得による経路決定法とQ-学習を用いた各AGV間の通信プロトコルの確立法を提案して、その有用性を検証した。これにより実工場で使用するための方向性を示すことができた。今後としては、デッドロックへの有効な対処法、各AGVの直交衝突の問題についての検討したい。

参考文献

- 1) PHILIPS D. WASSERRMANN : ニューラルコンピューティング-上級編-, 森北出版, (1998) 59~77
- 2) 加藤龍他 : ロボティクス・メカトロニクス講演会'99, 講演論文集, (1999)