

## 凹凸のある路面における足型ロボットの安定走行

函館高専 ○足達勇一 中村尚彦 浜 克己

## 要 旨

本研究では、凹凸のある路面における対象物の運搬作業を想定し、このときロボットがその状況に応じて一定の姿勢を保ちながら目標位置まで対象物を運搬することを目的として、凹凸の状態に柔軟に適應できる走行方法を確立するために、実機を含めた検討を行う。

## 1. 緒 言

近年のロボット産業の発展は目覚しく、使用される分野や用途も多様化している。それとともに、ロボットに要求されるタスクも複雑化し、あらかじめタスクに対して必要な知識を与えておくことが難しくなっている。そのため、ロボット自らが学び、考えて行動する自律学習型ロボットの要求が高まっている<sup>1)</sup>。

本研究では、凹凸のある路面における対象物の運搬を想定して、路面の状態に柔軟に適應できる学習型制御方式の構築を目指し、強化型学習の一つであるQ学習<sup>2)</sup>を用いて、行動を繰り返すことで最適経路の生成を試みる。また、実機による実現についても検討する。

## 2. 問題設定

本研究の問題設定を図1に示す。対象空間は3次元で管理されるものとし、ロボットが走行するときの路面に当たる高さ方向には凹凸があり、さらに固定ではあるが複数の障害物が存在し、それらも高さ情報を持つものとする。

本研究で使用するロボットの外観を図2に示す。ロボットは、上下機構を有する3本の脚（未装着）とその先に全方位移動が可能な車輪を持つ構造とし、移動の際して、姿勢を一定に保つ、すなわち運搬する対象物を載せる天板を水平に保持するような安定した走行の実現が要求される。実際には、平面的な移動方向と脚の伸縮という3次元的な行動が必要となり、障害物は水平・垂直の両方向で回避することになる。

ロボットは以下のように表現される。

$$R_L = \{P, D, J_i, S\} \quad (1)$$

$P$ : ロボットの位置,  $D$ : ロボットの進行方向

$J_i$ : 脚の伸縮状況 ( $i=1, 2, 3$ ),  $S$ : センサ情報

センサは、障害物検知・自己定位用に光学センサ、傾き検知用に加速度センサを使用する。

本研究での評価関数を以下のように定義する。

$$E = C_1 \cdot \sum_{i=0}^{step} S'_a + C_2 \cdot step \quad (2)$$

$S'_a$ : 時刻  $t$  における天板水平面からの傾き角度

$step$ : 目標位置に達するまでの総時間ステップ数

$C_1, C_2$ : 結合係数

これより、 $E$  を最小化するような経路を生成することが問題の解とする。

## 3. 行動選択

本研究では、ロボットの行動選択のためにQ学習を使用

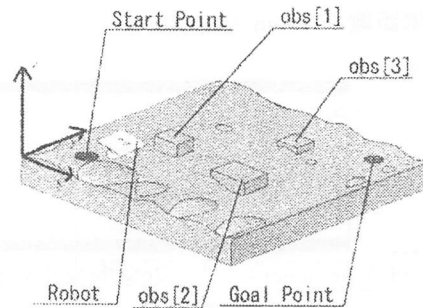


図1 問題設定

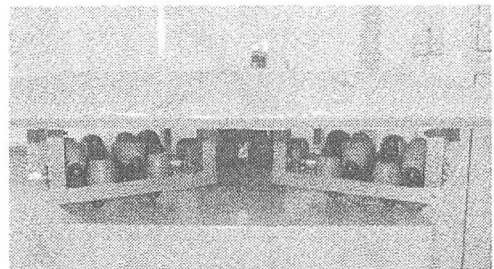


図2 ロボットの外観

する。Q学習は、与えられたタスクを達成するためにそれぞれの状態で行動をし、その結果として環境から報酬を受け取り、この報酬を最も多く獲得できるように各状態における行動の価値を更新していく。目的が経路探索の場合は、最終的に報酬を多く獲得した状態の並びが最適な経路として求められることになる。Q値の更新式を以下に示す。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} - \beta S'_a + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (3)$$

$s_t$ : 時刻  $t$  の状態,  $a_t$ : 時刻  $t$  の状態での行動

$s_{t+1}$ : 時刻  $t+1$  の状態,  $a$ : 時刻  $t+1$  の状態での行動

$\alpha$ : 学習率,  $\beta$ : 凹凸回避の重み,  $r$ : 報酬,  $\gamma$ : 割引率  
また、 $S'_a$  を Level gap として、以下の式を用いた。

$$S'_a = \left| \frac{h_{t+1} - h_t}{h_{\max}} \right| \quad (4)$$

$h_t$ : 時刻  $t$  の状態の高さ,  $h_{t+1}$ : 時刻  $t+1$  の状態の高さ

$h_{\max}$ : 環境内の最高点と最低点の差

本研究での状態  $s$  は、平面位置座標  $x, y$  と3つの脚の伸縮長さ  $J_i$  の組合せとなる。また、行動  $a$  は位置座標と脚の伸縮長さの各変化によって決定され、3次元的な動作とな

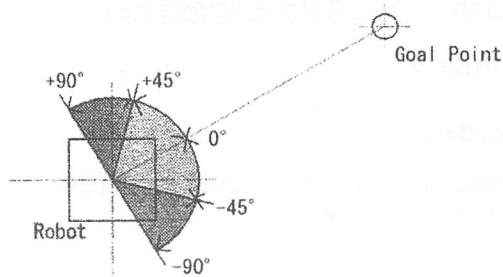


図3 行動選択

る。このとき、各時刻ステップでの天板の傾きは、常に一定範囲内となるのが条件である。そこで、従来の Q 値の更新式に、高さの変化をパラメータとして追加した。これにより、天板の傾きを小さくする行動の価値を高くする。

行動選択方法としては  $\epsilon$ -グリーディ手法を用いるが、Q 値が確率的な選択を行う場合、特に学習の初期段階では望ましい行動を選択することが困難である。対象物を搬送しながら目標位置に辿り着くことが目的であるため、行動選択に対して図3に示すように優先度を与えることとした。具体的には、目標位置の方向に対して  $\pm 90^\circ$  の範囲は  $10^\circ$  間隔に、さらに進行方向に対して  $\pm 45^\circ$  の範囲は  $5^\circ$  間隔にそれぞれ分割し、これらの方向を等確率で選択することとした。この設定により、目標位置方向への経路が選択される確率を上げるようにした。

#### 4. 計算機実験

ここでは、表面に凹凸があり、さらに通行不可能な障害物が存在する環境を設定し、ロボットが初期位置から目標位置までの移動経路を決定する問題に対して実験を行った。このとき、ロボットは上下機構を持たない構造とし、行動は水平方向のみの移動に限定した。Q 学習のパラメータとして、 $\alpha=0.1$ ,  $\gamma=0.9$  とし、さらに報酬は  $r=10$  とし目標位置のみに与えた。4種類の  $\beta$  の値について、ステップ数や総傾きの推移の調査、経路生成を実施した。

実験結果を図4から図6、および表1に示した。式(2)で定義した評価関数  $E$  について、結合係数を  $C_1=1.0$ ,  $C_2=15.0$  とし、 $S'_a$  には総傾き (Level gap) を用いた。この  $S'_a$  および Step は、25000~30000 エピソード間の最大値を用いた。図4、5より、ゴール時の Step がほぼ収束し、解として経路を生成できた。また、図4、6より、 $\beta=0.1$  のときに Step, Level gap ともに安定した。 $\beta=0.01$  でもほぼ同様の結果となったが、Episode が 25000 付近で  $\beta=0.1$  の時と比較して、Level gap が大きな値となっているため  $\beta=0.1$  の時のほうがよいと判断できる。表1より、評価関数  $E$  も  $\beta=0.1$  の時に最小となった。

本実験により、凹凸のある路面に対して、Q 学習を用いることで評価のよい経路が生成できた。

#### 5. 結言

本研究では、凹凸のある路面での安定した搬送作業を目的として、障害物回避を含め、それらに対応できるロボットの構造と制御方法について示した。

計算機実験によって、経路生成に対する学習方式を検討

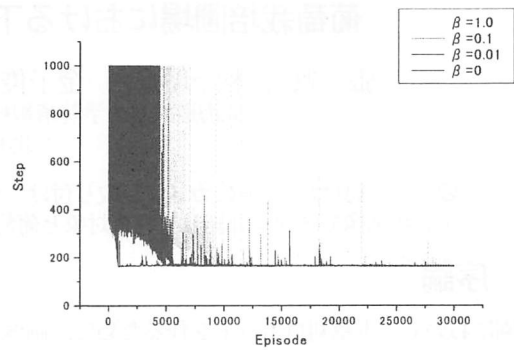


図4 エピソードによるゴール時のステップ推移(1)

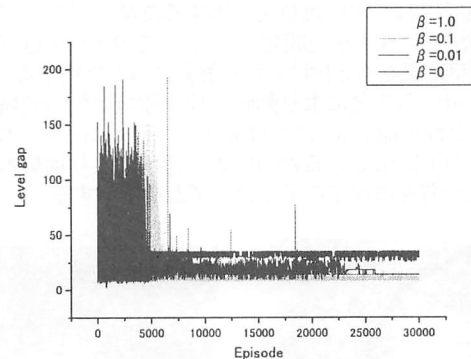


図5 エピソードによるゴール時の総傾き推移(2)

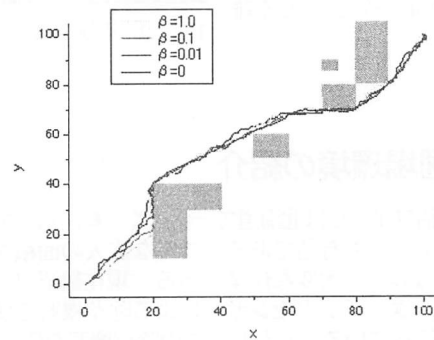


図6 生成経路

表1. 結果評価

	$\beta=0.0$	$\beta=0.01$	$\beta=0.1$	$\beta=1.0$
$E$	713.5522	438.5087	370.9825	565.1703

し、その有効性を確認することができた。今後は実機への実装に関して、検討が必要である。

#### 参考文献

- 1) 柴田克成, 強化学習とロボットの知能, 人工知能学会全国大会(第16回)講演論文集, 2A1-05, 2002
- 2) 高玉圭樹, マルチエージェント学習, コロナ社, 2003