2 移動ロボットシステムにおける知能と身体性の 構造的関連特性に関する研究*

澤田 務^{*1}, 須飼 淳 史^{*1}
市川 純 章^{*2}, 原 文 雄^{*3}

Structural Characteristic of Intelligence and Embodiment in Two-Mobile Robots System

Tsutomu SAWADA, Atsushi SUGAI,
Sumiaki ICHIKAWA and Fumio HARA^{*4}

^{*4} Department of Mechanical Engineering, Science University of Tokyo,
1-3 Kagurazaka, Shinjuku-ku Tokyo, 162-8601 Japan

This paper deals with an investigation to the relationship between robotic intelligence and embodiment. The robotic intelligence may be strongly related among the task the environment, and the body. In the task that two-mobile robots composed of crawler and 1 dof prismatic arm carry a long bar horizontally along a compound road with keeping the bar support length constant, the two pair of the robots, with different character of body (size of body and locomotion-distance per 1 step), acquire the intelligence autonomously in a simulation environment by reinforcement learning. We investigated the structural characteristic of relationship between robotic intelligence and the size of body, locomotion-distance per 1 step.

Key Words: Intelligent Equipment, Moving Robot, Learning, Embodiment, Simulation, Size of Body, Locomotion-Distance Per 1 step

1. 緒言

近年、自律知能ロボットの構築において、知能と同様に身体性も重要であることが徐々に認知されつつある⁽¹⁾⁽²⁾。身体性を扱った研究の中には、環境に働きかける身体を意識した自律知能ロボットの研究として、身体を持ったロボットが環境とインタラクションすることにより物体をクラスタリングする研究⁽³⁾や、人間型能動両眼視覚システムを用いて、内部結合をもたない首と眼の協調動作を、身体のみを経由して環境とインタラクションすることで学習させる研究⁽⁴⁾等がおこなわれている。また、身体が環境とインタラクションすることにより形態を形成する研究として、摩擦係数の異なる平面上での線状クラスタロボットによる形態形成に関する研究⁽⁵⁾や、ロボットと障害物との距離を推定するためのセンサ配置の形成に関する研究⁽⁶⁾等がおこな

われている。

ところで、知能という用語は様々な分野で用いられ、その定義は様々である⁽⁷⁾。本論文では知能を「センサ入力情報からアクチュエータ入力情報を決定する情報の変換則」と捉えている。

自律知能ロボットがタスク処理をおこなう際、「ロボットは、その身体に基づいて環境との動力学(運動学)的インタラクションをすることから、環境についての情報(センサ入力)を得て、その情報からアクチュエータを駆動するための情報を知能に基づいて決定する。そして、ロボットの知能、身体、動力学(運動学)を通して環境とインタラクションすることにより、与えられたタスク処理の能力が創発している(図1)」と筆者らは捉えている。つまり、自律知能ロボットのタスク処理において、与えられたタスクと環境の下で、ロボットの知能、身体、動力学(運動学)は独立に存在するものではなく、互いに強い関連があると考えられる。それらの内、設計者が関与できるのは知能と身体のみである。従って、与えられたタスクと環境において、ロボットの知能、と身体(材質、質量、大きさ、センサ、及びアクチュエータの種類、取り付け位置、

* 原稿受付 1999年12月16日。

^{*1} 東京理科大学大学院(〒162-8601 東京都新宿区神楽坂1-3)。

^{*2} 東京理科大学工学部。

^{*3} 正員、東京理科大学工学部。

E-mail: sawa@hafu.0103.me.kagu.sut.ac.jp

CPU、メモリ等)の関連を明らかにすることは、自律知能ロボットの研究で大切であるといえる。

Fig.1 Relationship among intelligence, body, dynamics or kinematics, environment and ability for task

知能と身体を同時に形成する方法として進化的手法等を用いた研究^{(8)~(12)}がなされているが、身体のどのような特徴が、知能とどのような関連を持つのかは十分に解明されていない。以上のような観点から、ロボットの知能と身体の特徴の関連特性を明らかにすることは、ロボット設計という分野において重要であるといえる。

知能と身体的设计に関して、トップダウン的な知能と身体的设计では、設計者の主観が入ることがあり、知能と身体の関連特性の評価には適さないと考えられる。そこで、様々な身体の特徴を持ったロボットに学習を用い自律的に知能を獲得させる。その知能はロボットの身体と環境とがインタラクションしながらロボットが自律的に獲得した知能であるので、設計者の主観が入り込むことがなく、客観的な評価ができると考えられる。

本研究では、平面と斜面で構成された複合路面という環境下で、前後進用駆動装置、伸縮方向に1自由度のアームを持つロボットモジュール二台が縦列して長尺物を支持幅一定、水平の拘束条件の下に運搬するというタスクを扱う。ロボットの身体の特徴の要素として「身体の大きさ」と「1ステップの移動量」を取り扱い、この2つのパラメータを変えた複数種類のロボットを用いて、学習による知能獲得のコンピュータシミュレーションを行う。平坦でない路面上で長尺物を支持幅一定、水平の拘束条件の下に運搬させるこのタスクではその協調行動が複雑となり、上記の身体の特徴が協調行動、すなわち獲得した知能の特性に与える影響が強いと考えられるので、この環境とタスクを採用する。

さて、センサ信号から張られるベクトル空間をセンサ信号と行動の対応の視点から離散的に分割したものを状態空間と定義する。ところで、先に述べたように本論文では知能を「センサ入力情報からアクチュエータ入力情報を決定する情報の変換則」と捉えているので、この獲得された知能は行動の視点からセンサベクトル空間を離散的に分割した状態空間で表現すること

ができる。そこで、異なる身体の特徴を持ったロボットに学習法 STL⁽¹³⁾を用い上述の状態空間を獲得させる。その状態空間の構造の特性にロボットの身体の特徴(身体の大きさと1ステップの移動量)が与える影響を解析することで知能と身体の構造的関連特性を明らかにする。

2. 知能獲得支援

2.1 STL 本研究ではロボットが自律的に知能を獲得する手法として、強化学習の一手法である STL⁽¹³⁾を用いる。STLについてその概要を以下に示しておく。

- ・ロボットは二つの行動モード(学習モード、タスク達成モード)を持ち、学習が十分できていない環境では学習モードに、学習が十分できている環境ではタスク達成モードに行動モードを自律的に遷移させる。
- ・学習モードにおいては、ロボットの行動の視点から同等と認識される状態を、センサデータを元に自律的に分割し状態空間を構成する。

従って、STLはロボットに自律的にセンサ入力情報からアクチュエータ入力情報を決定する情報の変換則、すなわち知能を獲得させるのに適した学習法であるといえる。以下にその概要を示す。

2.2 強化学習 STLの基本アルゴリズムは強化学習の一手法であるQ学習⁽¹⁴⁾である。時刻*t*での状態を*s*とし、ある行動*a*を選択した結果、時刻*t+1*で状態が*s'*に遷移し即座の報酬*r*を獲得したとき行動価値関数*Q_t(s,a)*を式(1)により更新する。ただし、*a'*は行動群*A*に含まれるある一つの行動を意味する。

$$Q_{t+1}(s,a) = (1-\alpha)Q_t(s,a) + \alpha \left\{ r + \gamma \max_{a' \in A} Q_t(s',a') \right\} \quad \dots\dots\dots (1)$$

α : Learning rate, γ : Discount rate

r : Immediate reward, A : Action set

そして、時刻*t*における行動選択確率*P_t(s,a)*は*Q*値を用いて式(2)で決定する。

$$P_t(s,a) = \frac{\exp(Q_t(s,a)/T)}{\sum_{a' \in A} \exp(Q_t(s,a')/T)} \quad \dots\dots\dots (2)$$

T : Boltzmann temperature

2.3 状態分割 ロボットは与えられた環境下で行動することにより、状態と行動の各組(*s,a*)について、ある設定された階級の幅(*R_t*)に従い報酬値とその獲得頻度の関係を得ることができるとする。ロボットがその環境下で行動を続け、上記の関係を構築する過程で、ある(*s,a*)において二種類の報酬値(*r₀*, *r₁*: $|r_0| < |r_1|$)が報酬値の獲得頻度の予め設定されたしきい値(*T_f*)を越えた

とする。このとき、もともと認識していた状態 s において同じ行動 a を選択しても獲得する報酬に違いがあるのは異なった状態を1つの状態として誤認識していたためとし、これを状態分割のきっかけとする。

状態分割のきっかけを得たロボットは、状態を構成する要素であるクラスをセンサ次元ごとにつくる。分割される状態は、各センサ次元でのクラスの組み合わせで表現される。ロボットは与えられた環境下で行動することにより、各 (s, a) においてセンサ次元ごとにセンサデータと報酬値の関係を得ていたとする。ここで報酬の獲得頻度が T_f を越えた二種類の報酬値のうち、絶対値の大きい報酬値 (r_1) に関するセンサデータのみを用いてクラスを構成する。クラス構成の方法として隣り合う各センサデータ値の差が式(3)に示す条件を満たす場合、それらのセンサデータはひとつのクラスとして扱うことにする。つまり、行動 a をとることによって、より大きい報酬値の獲得できるセンサデータの範囲をクラスとする。そして、センサ次元ごとのそれらクラスの組み合わせを新しい状態として状態 s から分割する。

$$D_i - D_{i-1} \leq T_d \quad (i: 1 \sim n) \quad \dots\dots\dots (3)$$

$$T_d = \frac{R_d (D_{\max} - D_{\min})}{\sqrt{n}} \quad \dots\dots\dots (4)$$

T_d : Threshold of disparity, R_d : Rate of disparity

$D_{\max}$: Maximum data, $D_{\min}$: Minimum data

D_i : Sensory data for r_1 , n : Sensory data number for r_1

2.4 行動モードの自律的切り替え Q 値に従ってどの程度学習結果を行動に反映させるかを決定するパラメータは、行動選択確率を決定する式(2)におけるボルツマン温度である。STL では、この温度をロボットが自律的に変化させることにより、ロボットは行動モードを学習モードとタスク達成モードの間で自律的に遷移させる。

温度が高いとき、行動選択確率は際立ったピークを持たないので、ロボットの行動は学習的となる。また、温度が低いとき、ロボットは各状態において最も高い Q 値に従った行動を選択するので、ロボットの行動はタスク達成的となる。そこで、この温度をロボットの行動モードを示すパラメータ " T_s " とし、ロボットの獲得する報酬の履歴に基づいて更新する (式(5))。

$$T_s(t+1) = (1 - T_{cr})T_s(t) + T_{cr}T_m \quad \dots\dots\dots (5)$$

$$\text{if reward} = 0, \text{ then } \begin{cases} T_{cr} = T_{cr \max} \\ T_m = T_{\max} \end{cases}, \quad \text{else } \begin{cases} T_{cr} = T_{cr \min} \\ T_m = T_{\min} \end{cases}$$

$T_{\max}$: Maximum T_s , $T_{cr \max}$: Convergence-rate to $T_{\max}$

$T_{\min}$: Minimum T_s , $T_{cr \min}$: Convergence-rate to $T_{\min}$

Q 学習に基づくロボットの行動選択はボルツマン温度に依存して確率的に変動する。また、式(5)で決まる

T_s には、学習モードとタスク達成モードを切り替えるしきい値はなく、どの行動モードが発現するかの度合いは、 T_s によって表現される。ロボットが獲得する報酬の履歴により、ロボットは、自律的に行動モードを、学習が十分できておらず報酬が十分獲得できない環境では学習モードに、学習が十分できていて報酬が十分獲得できる環境ではタスク達成モードに遷移させる。

3. シミュレーション設定

3.1 環境とタスク ロボットが移動する環境は図2に示す一次元複合路面とする。各ロボットは前後進用の駆動装置、及び伸縮方向に1自由度のアームを持つ。ロボットに与えるタスクは縦列する二台のロボットモジュールで長尺物を支持幅一定(600 [mm])、水平(0 [deg])の拘束条件の下に運搬することとする。ただしこのとき、支持幅一定、水平にはある許容範囲を与える。

Fig.2 One dimensional compound road

3.2 行動空間 ロボットの行動はそれぞれアームの静止、上昇、下降、及び駆動装置の前進、後進の組み合わせからなり、その行動空間 $A = \{0, 1, \dots, 8\}$ を図3の様に設定する。

Fig.3 9 actions constituting action set A

3.3 初期状態空間と初期 Q 値 二台のロボットモジュールが運搬する長尺物の支持幅 (sl) と支持傾き (θ_{sl}) を二次元センサデータとする。状態空間は (sl, θ_{sl}) の二次元平面上で表現することができる。初期状態空

間を、支持幅一定、且つ水平の許容範囲内($580 \leq sl \leq 620$ 且つ、 $-1.5 \leq \theta_{sl} \leq 1.5$)を $s=1$ として設定し、他は全て $s=0$ とする。初期 Q 値として $Q(1,0)=1$ を設定し、他の Q 値は全て 0 とする。これは支持幅一定、且つ水平の許容範囲内を示す状態($s=1$)では、二台のロボットモジュールがアームを静止させたまま前進($a=0$)することに対する行動価値を高くしておくことを意味する。(図 4)

Fig.4 Initial state space and initial Q value

3.4 報酬 長尺物の支持状態が支持幅一定、且つ水平の許容範囲を保ったまま二台のロボットモジュールが前進した場合は、即座の報酬 $r=1$ をロボットに与える。

長尺物の長さは有限であるため、学習モードにおいて長尺物を落とす危険がある。そこで支持可能範囲を図 4 のように設定し、これを越えた場合、即座の報酬 $r=-2$ をロボットに与える。そして、次の時刻での行動は式(2)に従って確率的に選択するのではなく、無条件に元に戻ろうとする行動を選択することにする。それ以外は即座の報酬 $r=0$ とする。

ロボットに与えられた行動空間は図 3 に示すように(後行ロボットモジュール, 先行ロボットモジュール)に対して、アームの(静止, 静止)(下降, 上昇)(上昇, 下降)、駆動装置の(前進, 前進)(後進, 前進)(前進, 後進)の組み合わせで構成されている。図 2 に示す路面右端まで長尺物を運搬するというタスクの観点から、駆動装置の(前進, 前進)は(後進, 前進)(前進, 後進)より少ないステップ数でタスクを達成することができる。よって、駆動装置のステップ数の大小を消費エネルギーの大小で表現する。一方、アームの(静止, 静止)は(下降, 上昇)(上昇, 下降)より、消費エネルギーが小さいことは明らかである。以上の観点から、アーム、駆動装置の行動に対する消費エネルギーを決める係数をそれぞれ Ec_{arm} 、 Ec_{drv}

と定義する。そして Q 値を式(6)に基づいて更新することにより、行動の消費エネルギーを報酬に反映させる。

$$Q_{i+1}(s, a) = (1 - \alpha)Q_i(s, a) + \alpha Ec_{arm} Ec_{drv} \{ r + \gamma \max_{a' \in A} Q_i(s', a') \} \quad \dots\dots\dots(6)$$

Ec_{arm} : Energy-consumption rate for arm

Ec_{drv} : Energy-consumption rate for drive

3.5 身体の特徴 図 5 に示す身体の大きさ(Size of body: Sb)と 1 ステップの移動量(Locomotion distance per 1step: Ld)の異なる十六種類のロボットを扱う。前後のロボットは同じ身体の特徴を有している。なお括弧内の数値はそれぞれ (Sb - Ld) を表す。

Fig.5 16 robots composed of different characteristics of body

3.6 各パラメータ 表 1 に学習におけるパラメータを示す。また、ロボットの初期状態は $sl=600$, $\theta_{sl}=0.0$, $T_s=T_{min}$ とする。これらの値を用いて、学習シミュレーションを実施した。その他実験の設定の詳細は文献(13)に従うものとする。

Table 1 Parameters for Learning

α	0.1	γ	0.8
T_f	20	R_d	0.5
R_r	0.1		
T_{min}	0.01		
$T_{cr min}$	0.8	T_{max}	5.00
		$T_{cr max}$	0.1
Ec_{arm}	If $a = 0, 3, 6$ then $Ec_{arm}=1.0$ Else then $Ec_{arm}=0.6$		If $a = 0, 1, 2$ then $Ec_{drv}=1.0$ Else then $Ec_{drv}=0.6$

4. 実験

3章に示した設定に従って、身体の特徴の異なるロボットに学習をおこなわせる。ロボットが路面右端のゴールに到達するまでを一試行とし、試行を繰り返すおこなうことにより学習を進める。本研究で用いる学習法 STL では、学習モードとタスク達成モードの度合いが T_s によって表現され、自律的に行動モードが遷移する。従ってある試行において、行動モードを示すパラメータ T_s が全てのステップで $T_s = T_{min}$ であったとき、すなわち行動モードが十分にタスク達成モードであったとき、学習が収束したと判断する。

各ケースにおいて行動選択における乱数の種を変えて、長尺物を運搬する知能を 50 個獲得させる。

5. 結果、及び考察

5.1 獲得された知能 図 6 に、ケース 15 ($Sb=200$, $Ld=10$)におけるセンサ状態を各行動の視点から分割した状態空間、及び二次元センサデータ支持幅 sl 、支持傾き θ_{si} の軌跡の例を示す。

本論文では知能を「センサ入力情報からアクチュエータ入力情報を決定する情報の変換則」と捉えている。従って、知能を行動の視点から離散的に分割された状態空間で表現することができる。ロボットが学習によって獲得する知能は、個々の経験（環境とのインタラクション）に基づくものである。従って、ロボットの行動が確率的に選択される学習法 STL では、同じ身体の特徴を持ったロボットにおいても、行動選択における乱数の種が異なれば、同じ知能を獲得するとは限らない。例えば、図 6(a), (b) はセンサデータの軌跡は同じ、すなわち、各ステップにおいて選択した行動の時系列は同じであるにも関わらず、状態空間という視点からは異なるものである。

Fig.6 State space partition by action and locus of 2 dimensional sensory data in case 15

そこで、次節以降に身体の特徴が知能にどの様に影響するかを調べるために、各行動の使用率、各行動が選択されるセンサ状態の傾向の観点から知能構造を比較し、知能と身体の大きさ、及び 1 ステップの移動

量の関連特性を明らかにする。

5.2 知能と身体の大きさの関連特性 図 7 に、各ケースにおいてロボットが長尺物を運搬する際の各行動の使用率を示す。例えば、使用率 0.8 は 50 個の知能の内、40 個の知能が、長尺物運搬時に少なくとも一回以上その行動を使用することを意味する。

この結果から、1 ステップの移動量 Ld に関わらず、身体の大きさ $Sb=50, 100$ のロボットでは、 $a=0, 1, 2$ の行動の使用率が高く、 $Sb=150, 200$ のロボットでは、 $a=0, 1, 2, 3, 6$ の行動の使用率が高いことがわかる。

実際には、身体の大きさ $Sb=50$ のケース 1~4 ではそれぞれ 54%, 68%, 82%, 100% の知能が、 $Sb=100$ のケース 5~8 ではそれぞれ 52%, 70%, 86%, 86% の知能が $a=0, 1, 2$ の三種類の行動のみを使用し長尺物を運搬する知能であり、身体の大きさ $Sb=150$ のケース 9~12 ではそれぞれ 68%, 58%, 68%, 96% の知能が、 $Sb=200$ のケース 13~16 ではそれぞれ 58%, 54%, 62%, 92% の知能が $a=0, 1, 2, 3, 6$ の五種類の行動のみを使用し長尺物を運搬する知能である。

Fig.7 Rate of use for each action in each case

図 8 に、与えられたタスクに対して各身体の大きさを持つロボットが、 $a=0$ 、すなわち、アームを静止させたまま前進した際の支持幅 sl と支持傾き θ_{si} の二次元センサデータの軌跡を示す。そして、二次元センサデータの軌跡の観点から知能と身体の大きさの関連特性を考察する。

支持傾き θ_{si} が $\theta_{si}=0$ [deg] からずれるときの最大値は身体の大きさに関わらず一定であるが、支持幅 sl が $sl=600$ [mm] からずれるときの最大値は身体の大きさに依存し、身体の大きさが大きいほどずれの最大値は大きい。

$Sb=50, 100$ のロボットにおいては、 θ_{si} が水平とみな

される範囲を超えることはあるが、 sl が支持幅一定とみなされる範囲を超えることがないため、支持幅を補正する $a=3\sim 8$ を使用しないで長尺物を運搬する知能を獲得することが可能である。一方、 $Sb=150, 200$ のロボットは θ_{si} に加え、 sl も支持幅一定とみなされる範囲を超えることがあるため、 $a=0\sim 2$ に加えて支持幅を補正する $a=3\sim 8$ 、その中でも消費エネルギーの少ない $a=3, 6$ を使用し長尺物を運搬する知能を獲得する傾向にある。

図 8 に示す二次元センサデータの軌跡は、身体の大きさのみに依存し、1ステップの移動量には依存しないので、 $a=0, 1, 2$ の三種類の行動を使用して長尺物を運搬する知能か、 $a=0, 1, 2, 3, 6$ の五種類の行動を使用して長尺物を運搬する知能かは身体の大きさのみで決まる。その際、身体の大きさが $Sb=120$ のとき、支持幅 sl のずれの最大値が支持幅一定とみなされる範囲と一致する。従って、 $a=0, 1, 2$ の三種類の行動を使用して長尺物を運搬する知能、 $a=0, 1, 2, 3, 6$ の五種類の行動を使用して長尺物を運搬する知能のいずれかを獲得するかは身体の大きさの臨界値 $Sb=120$ によって決まるといえる。身体の大きさの臨界値は環境や、タスクの条件によって決まる値である。

以上の結果から、同一の環境での同一タスクに対して、身体の大きさの異なるロボットは、長尺物運搬の際に必要な行動が異なる知能を獲得することがわかる。

Fig.8 Locus of 2 dimensional sensory data for each Sb

5.3 知能と移動量の関連特性 ロボットが長尺物を運搬する際、16 ケースの各々において、各行動がどのようなセンサ状態で選択される傾向にあるのかを、マハラノビス距離を用いた式(7)に示す二次元楕円体で表現⁽¹⁵⁾する。ただし、 $a=0$ を除く、前節で述べた行動の使用率が高い (0.8 以上) 行動に関してのみを扱う。

$$(x-\mu)^T \Sigma^{-1} (x-\mu) = m+2 \quad \dots\dots\dots (7)$$

$x: (sl, \theta_{si})$ μ : The mean vector of x

Σ : Variance covariance matrix of x

m : Sensory dimension ($m=2$)

その結果を図 9 に示す。この結果から、1ステップの移動量の大きいロボットは各行動を選択するセンサ状態の領域が小さいことがわかる。

図 10 に、身体の大きさ Sb の各々において、各行動に対応する楕円の平均半径 $\sqrt{s}$ を示す。ただし、楕円の

面積 S を算出する際、各センサ次元をそれぞれ、支持幅一定、支持傾き水平とみなされる範囲の大きさと無次元化する。すなわち、支持幅一定、支持傾き水平とみなされる範囲の面積の大きさは $S=1$ となる。

Fig.9 Ellipse of state partition for each case

Fig.10 Mean radius of ellipse for each action in each Sb

本論文において、知能ロボットシステムにおける知能とは、センサ入力情報からアクチュエータ出力情報を決定する情報の変換則であることから、制御系における制御器と捉えることができる。そこで、制御系の観点から、知能と移動量の関連特性を考察する。

本研究で用いたロボットシステムは、その移動量が固定されるとき、離散システムと考えることができる。このロボットシステムは学習シミュレーションにおいて運動学的移動をしているが、長尺物の支持状態が支持幅と支持傾きの許容範囲に入ることによって「安定」とみなし、これを一般化して動力学的に考察することにする。

一般的に力学系は二次遅れ系で表現することができるので、二次遅れ系 $G(s)$ (式(8)) を制御対象とし、フィードバック比例制御をおこなうサンプル値制御系を考える (図 11)。ただし、ホールド回路 $G_H(s)$ は 0 次ホールド回路 (式(9)) とする。

$$G(s) = \frac{K}{s(s+1)} \dots\dots\dots (8)$$

K : Gain constant

$$G_H(s) = \frac{1-e^{-Ts}}{s} \dots\dots\dots (9)$$

T : Sampling period

Fig.11 Sampled data control system

このとき、パルス伝達関数 $G(z)$ は

$$\begin{aligned} G(z) &= \mathcal{Z}[G_H(s)G(s)] = (1-z^{-1})\mathcal{Z}\left[\frac{G(s)}{s}\right] = (1-z^{-1})\mathcal{Z}\left[\frac{K}{s^2(s+1)}\right] \\ &= K \frac{(T+e^{-T}-1)z + (1-e^{-T}-Te^{-T})}{(z-1)(z-e^{-T})} \dots\dots\dots (10) \end{aligned}$$

となる。

本研究で用いたロボットシステムがタスクを遂行している状態は、制御系の観点から安定状態にあると考えられるので、その安定性を検討する。この制御系の特性方程式は $1+G(z)=0$ で与えられる⁽¹⁶⁾。従って、

$$\begin{aligned} 1 + K \frac{(T+e^{-T}-1)z + (1-e^{-T}-Te^{-T})}{(z-1)(z-e^{-T})} &= 0 \\ z^2 + \{K(T+e^{-T}-1) - (1+e^{-T})\}z + K(1-e^{-T}-Te^{-T}) + e^{-T} &= 0 \end{aligned} \dots\dots\dots (11)$$

以上より、Jury の安定判別法⁽¹⁵⁾による安定の条件は

$$D(z) = a_0z^2 + a_1z + a_2 \dots\dots\dots (12)$$

$$a_0 = 1$$

$$a_1 = K(T+e^{-T}-1) - (1+e^{-T})$$

$$a_2 = K(1-e^{-T}-Te^{-T}) + e^{-T}$$

において、 $|a_2| < a_0$, $D(1) > 0$, $D(-1) > 0$ であるから、

$$\frac{-(1+e^{-T})}{1-e^{-T}-Te^{-T}} < K < \frac{1-e^{-T}}{1-e^{-T}-Te^{-T}}$$

$$K > 0$$

$$K < \frac{2(1+e^{-T})}{-2+T+2e^{-T}+Te^{-T}}$$

を満たすことである。図 12 に K と T に関して安定範囲を示す。図 12 は、ゲイン K が大きくてもなお、システムが安定であるためにはサンプリング周期 T が小さくなければならないことを示している。

一方、ゲイン K はシステムの状態変化の大きさを決

定する係数であり、サンプリング周期 T は状態を認識する時間幅に対応している。 T が大きくなると状態の変化分が大きくなる可能性がある。つまり、これら 2 つの量がある許容範囲にあるとき、「安定」とであると解釈できる。そこで、長尺物の支持幅や支持傾きの変化の大きさがある許容範囲に入っていることと、この「安定」とが対応すると考える。従って、学習が収束した知能を持つシステムはこのような観点から「安定」とあると考えられる。

このことは、図 12 のゲイン K とサンプリング周期 T の安定範囲の境界線が図 10 の L_d と $\sqrt{s}$ の関係と類似していることに対応していると考えられる。

Fig.12 Stable range

6. 結言

本研究では、前後進用駆動装置、伸縮方向に一自由度のアームを持ち、縦列するロボットモジュール二台からなるロボットに、平面と斜面で構成された複合路面上で長尺物を支持幅一定、水平の拘束条件の下に運搬させるタスクを与えた。ロボットの身体の特徴として、「身体の大きさ」と「1 ステップの移動量」を変えた複数種類のロボットが、学習によって獲得する知能について、その構造を比較し、知能と身体の関連を解明した。

その結果、身体の大きさの異なるロボットは、同じ環境、同じタスクにおいても身体の大きさの違いにより環境から受けるセンサデータと環境への行動が異なるため、知能構造において、長尺物運搬の際に必要なとする行動の種類が異なることがわかった。本ロボットシステムの場合、身体の大きさが $Sb \leq 120$ のとき、 $a=0, 1, 2$ の三種類の行動を使用して長尺物を運搬する知能、 $Sb > 120$ のとき $a=0, 1, 2, 3, 6$ の五種類の行動を使用して長尺物を運搬する知能を獲得する傾向があることがわかった。

一方、1 ステップの移動量の異なるロボットは、知能構造において、それぞれの行動を選択するセンサデータ領域の大きさが異なることがわかった。そして、本ロボットシステムと一般的なサンプル値制御系との比較により、本ロボットシステムにおける 1 ステップ

の移動量、及び状態変化の許容の大きさを示す楕円の平均半径は、離散システムの安定条件におけるゲインとサンプリング周期に定性的に対応していると考察されたが、より厳密で定量的な議論は今後の研究課題の一つである。

本研究では、身体の特徴として、身体の大きさと、1ステップの移動量を扱い、それと獲得された知能との構造的関連特性の解明をおこなったが、他の身体の特徴(材質、質量、センサ、及びアクチュエータの種類、取り付け位置)と知能の構造的関連特性の解明が他の研究課題である。

参考文献

- (1) 浅田稔、身体性による知能の発現、人工知能学会誌、13-1 (1998)、14-15.
- (2) 國吉康夫・ベルトゥーズ リュク、身体性に基づく相互作用の創発に向けて、日本ロボット学会誌、17-1 (1999)、29-33.
- (3) H. Nakamura, A. Ishiguro, Y. Uchikawa and R. Pfeifer, EXPLOITING EMBODIMENT TO SOLVE SOME OF THE HARD CLASSIFICATION PROBLEMS: AN EMBODIED COGNITIVE SCIENCE APPROACH, Proceedings of International Workshop on Signal Processing Applications & Technology, (1998), 163-177.
- (4) L. BERTHOUSE and Y. KUNIYOSHI, Emergence and Categorization of Coordinated Visual Behavior Through Embodied Interaction, Machine Learning 31 (1/2/3), (1998), 187-200.
- (5) N. Kawai and F. Hara, Formation of Morphology and Morho-Function in Liner-Cluster Robotic System, Proceedings of the 5th International Conference on Simulation of Adaptive Behavior, (1998), 459-464.
- (6) L. Lichtensteiger and P. Eggenberger, Evolving the Morphology of a Compound Eye on a Robot, Proceedings of the 3rd European Workshop on Advanced Mobile Robots, (1999).
- (7) R. Pfeifer, C. Scheier, Understanding Intelligence, MIT Press.
- (8) C. W. Reynolds, An Evolved, Vision-Based Behavioral Model of Coordinated Group Motion, Proceedings of the 2nd International Conference on Simulation of Adaptive Behavior, (1992), 384-392.
- (9) K. Sims, Evolving 3D Morphology and Behavior by Competition, Artificial Life IV Proceedings, MIT Press, (1994), 28-39.
- (10) W. P. Lee, J. Hallam, H. H. Lund, A Hybrid GP/GA Approach for Co-evolving Controllers and Robot Bodies to Achieve Fitness-Specified Tasks, Proceedings of IEEE 3rd International Conference on Evolutionary Computation, (1996), 384-389.
- (11) H. H. Lund, J. Hallam and W. P. Lee, Evolving Robot Morphology, Proceedings of IEEE 4th International Conference on Evolutionary Computation, (1997)
- (12) 菊池耕生・原文雄、自律機械システムの構造形態と知能の進化的設計法に関する研究(第2報、汎化性と形態および知能構造の特性に関する一考察)、日本機械学会論文集(C編)、65-638 (1999)、150-157.
- (13) 澤田務・市川純章・原文雄、2 移動ロボットシステムのオンライン学習における行動モードの自律的切り替え、日本機械学会論文集(C編)、65-635 (1999)、202-209.
- (14) C. J. C. H. Watkins, Learning from delayed rewards, Ph. D. thesis, King's College, University of Cambridge, (1989).
- (15) 浅田稔・野田彰一・細田耕、ロボットの行動獲得のための状態空間の自律的構成、日本ロボット学会誌、15-6 (1997)、886-892.
- (16) E. I. Jury, SAMPLED-DATA CONTROL SYSTEMS, John Wiley and Sons Inc., (1958).