

自律型海中ロボットの訓練システム

—その1：障害物回避行動の教示と自己学習—

Training System for Autonomous Underwater Vehicles

—First Report: Teaching and Self-Learning for Collision Avoidance—

川 野 洋*・浦 環*

Hiroshi KAWANO and Tamaki URA

1. はじめに

人間が立入不可能な危険海域において、広範囲にわたる海洋調査を自動で行うことが出来る自律型海中ロボット (Autonomous Underwater Vehicle: AUV) の研究が行われている。AUVを使用するのは、ソフトウェア工学の専門家ではなく、海洋科学の専門家である。よって、ロボットの専門外の人でも簡単にAUVの行動制御アルゴリズムを作成でき、かつ作成されたアルゴリズムの信頼性を簡単に確保する機能を持つシステムが必要となる。本研究では、そのような「自律型海中ロボットの訓練システム」について「潮流中での障害物回避による目的位置への直線航行の訓練」を例にとって提案する。

2. 教示と自己学習

複雑な任務行動を行なうAUVのアルゴリズムには、信頼性が求められる。そのような行動をAUVに学習させるには、人が積極的に教示を行なう必要がある。ところが、行動内容の全てを教示する方式は教示者に多くの負担を強いるので、その負担が自己学習によって軽減できれば便利である。また、自己学習によって獲得された行動は、教示者の予想を越えた良いものになる可能性がある。以上の理由から、本研究の扱う訓練システムは、自己学習と教示を併用した方式とする。

続いて、ロボットに自己学習させる行動と教示を行なう行動の切り分けの方法について考える。ある行動がロボットにとって自己学習可能である条件とは、「行動と結果の写像関係が、ロボットの認識可能なパラメータを使った決定性計算により記術可能」であることである。そのような行動については、ロボットの認識するパラメータにより、一つの行動に対する結果が一つに定まるので、ロボット単体で適切な行動モデルが取得できる。以後これを、自己学習可能条件と呼び、これを満たす行動についてはロボットに自己学習をさせ、そうでない行動をロボットに教示することとする。一般にロボ

ット的环境認識能力は限られているので、高度な行動を自己学習させる事は困難であるが、教示によって、そのようなロボットの学習限界を引き上げることがねらいである。

具体的には、想定任務を行なうAUVが超音波レンジセンサー値、機体の対地速度値、潮流の速度値 (6自由度)、スラスト出力値を認識可能な場合、潮流速度値が決まれば、スラスト出力値、機体速度値の静的なつりあい関係は決定性計算で表現可能であるので自己学習可能である。逆に、超音波レンジファインダーの値と機体がとるべき速度値の関係は、自機位置と地形形状を表現したパラメータ抜きには、決定性計算で表現することは不可能であるので教示が必要である。

3. 自己学習の方法

自己学習により学習される行動モデルを、経験モデル (experienced model: EXM) とする。ロボットが任務中に未知の状況に出会うことを予防するためと、教示者の予想を越えた適切な行動をロボットがとることを可能にするためには、自己学習可能な事象をEXMにもれなく学習させる機能が必要である。例えば、潮流中でのAUVの挙動は複雑であり簡単に予測できないが、自己学習によって、存在しうる潮流下のスラスト値と速度値の写像関係がもれなくEXMに学習されていれば、どのような潮流中でも適切なスラスト値を計算できるし、教示者の予測を超えた行動で海中を泳ぐことができる。例えば、機体に対して横向きの潮流があるときに、AUVが単純に前進しようとするれば斜めに流されるが、その関係をEXMが保持していれば、斜め方向への航行指示があった場合に潮流に流されながら斜めに航行していく行動をロボットがとることが出来る。

このための方法として、ロボットの状態の集合をA、行動の集合をB、行動の結果の集合をCとして、AとBのCartesian積A×Bの要素とCの要素間の写像関係

$$f: A \times B \rightarrow C \dots\dots\dots (1)$$

を測定し、それから行動モデルの写像

*東京大学生産技術研究所 海中工学研究センター

$g: A \times C \rightarrow B \dots \dots \dots (2)$
 を作ることにする。ここで任意の $A \times C$ の要素 (入力) に対して、必ず B の値 (出力) が存在することは保証されていない。これを EXM を使用する時に考慮しなければならないのは不便である。

例えば、AUV には運動能力に制限があるので、潮流の強さと方向によっては、実現不可能な航行速度が存在する。しかし、EXM に対して速度指令を出す側で、そのような速度の制限を全て考慮するのは難しい。

よって、EXM には対応する出力値を持たない入力値を受け取っても、類推によって望ましい出力値を出す機能を持たせることにする。

4. 教 示 の 方 法

4.1 教示用のインターフェース

教示による行動モデルを訓練モデル (Trained Model: TM) と呼ぶことにする。信頼性のあるアルゴリズムの教示を軽い負荷で行えるためには、以下の機能が必要である。

- ① ロボットにいろいろな技能を独立に教示した後、それらを組み合わせて行動アルゴリズムが簡単に作成出来る機能。
- ② ロボットに教示された内容が、ロボットにどのように学習されているかを、教示者が教示作業中に容易に確認できる機能。
- ③ 教示のプロセスが、基本的な内容から始まり、段階を追って内容が詳細化していく構成であること。
- ④ 始めから教示内容の詳細が決まっても、教示作業を進めながら方針の詳細を詰めていくことが出来る機能。
- ⑤ ロボットが任務中に未知の事象に出会わないように、教示漏れを防止する機能。

以上から、ロボットに対して教示する内容を、

- ① センサー観測値の抽象化の仕方
- ② センサー抽象化表現と行動の間の写像関係

とし、①については、粗い解釈から段階的に細かい解釈を教示していく方法とし、それをセンサー観測値のなす状態空間を分割していくことにより行なうことにする。②については、「教示者がロボットを動かしてみせる形で教示入力を与え、ロボットは教示を受けたその場で真似事のような形で教示された行動をやって見せ、それに対して教示者が修正教示を加える」形の教示手法をとることにする。

4.2 学習則

学習済みの TM の信頼性を確認しやすくするために、TM の表現形式として有限状態機械¹⁾を採用する。学習則

として、ニューラルネットワークのように教示頻度が高い写像関係を学習するしくみを有限状態機械用に構築すれば、教示者がたまに誤った内容の教示をしても、正しい内容を学習させることが出来る。そこで、有限状態機械の状態をロボットの行動、アルファベットのセンサー観測値の抽象化表現とし、「各状態間をつなぐアルファベットとして、教示入力回数 of の最も多かったものを採用する」方法をとった。

5. アルゴリズムの信頼性確保の方法

教示者がロボットの行動の安全性について整理をしやすいように、行動の安全管理と意思決定を分離して考えられるしくみが必要である。また、間違った内容の教示を防止する機能が必要である。そのために、ロボットに禁止された行動を算出するための禁止行動計算モデル (Forbidden Command Calculating Model: FCCM) を行動モデルとは独立に導入し、FCCM によって教示作業と任務行動を監視する。

6. 行動モデルの構成

6.1 全体構成

行動モデルは図 1 のような構成となり、EXM の上流側に TM を接合する。EXM, TM の内部基本構造は Brooks の階層構造²⁾とし、各々の技能内容 (障害物回避, 目的位置への航行等) を各階層のモジュールに学習させる。通常、階層構造においては、各層モジュールはいつ出力をしないかについての情報を保持している。それは別階層のモジュールとの関係を考慮した情報であり、任務内容に依存する。そこで、各モジュールがその技能内容によって真に独立性を保つことができるように、各層共通に FCCM を用意し、各層の出力値がロボットにとっての禁止行動かどうかを監視させる。もし、禁止行動と各層の出力が一致している場合は、FCCM はその層の出力値を通さないようにする。これにより、任意の技能モジュールを実行優先度の高い順に上層から配置すれば、行動モデルを完成させることが出来る。また、FCCM の導入により行動モデルに類推機能を持たせることが出来る。なぜなら、各層のモジュールが出力する値が教示者の意図を完全には反映してない場合、FCCM によってそれが危険ではない妥協的な行動なのかをロボットが判断できるからである。行動モデルの学習は、FCCM → EXM → TM の順に進められる。

6.2 禁止行動計算モデルの構成

FCCM は、表 2 のようにセンサー観測値に対する禁止行動の写像関係を定義したものである。超音波レンジセンサによる障害物回避航行をする AUV を例にとれば、短い距離を観測したレンジセンサの向きへの航行を禁止行動とす

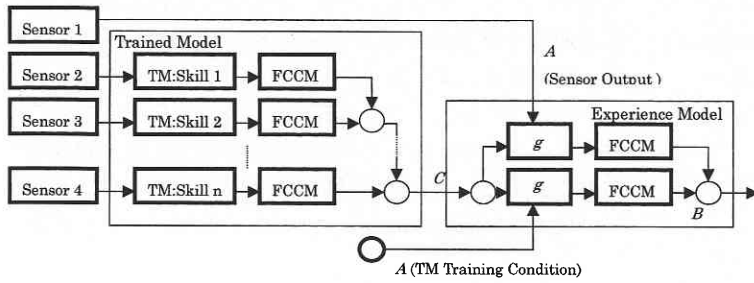


図1 行動モデルの構成

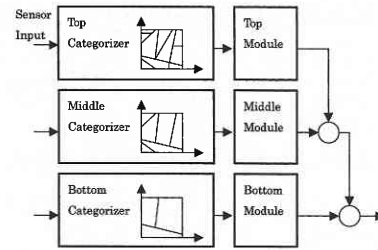


図2 各層技能モジュールの構成

ることで簡単に FCCM を教示できる。

6.3 自己学習モデルの構成

EXM を 2 層の階層構造とし、上下両層に学習済みの写像 g のモジュールを組み込む。上層に状態パラメータとしてセンサー観測値を入力し、下層には TM 訓練時の状態値を固定入力する。この下層のモジュールにより、本来ならば写像 g が対応する出力値を持たない入力値に対して、TM の教示時の状況から類推される妥協出力値を計算できる。

6.4 教示による行動モデルの構成

TM は階層構造をとり、階層毎に各技能のモジュールが置かれている。ロボットが任務中に想定外のセンサー入力値を受けるとを避けるために、各モジュールに入力される値のすべてに対して出力値を存在させる必要がある。そのためには、入力値の状態空間全体を適切に区分することによって抽象化し、区分されたすべての領域に対して出力値を割りあてることが必要である。しかし、区分の仕方が細かいほど多くの入力値の事例に対する出力値を定義しなければならず、教示作業が困難になる。そこで、各技能モジュールの内部構成を、抽象化の細かさの程度が異なる複数のモジュールからなる階層構造とした。図 2 中の Categorizer は、入力値を各階層のモジュールが解釈できる抽象表現に翻訳する。上層ほど抽象化表現が細かい。例えば、入力値が前方の地形についての情報を持っているとしたら、最下層の Categorizer では、地形は「上り坂、平地、下り坂」中間の層では、「きつい坂、緩い坂、山の頂上、緩いくだり……」を出力する。この構成により、下層のモジュールで教示漏れを防ぎながら、上層のモジュールで詳細な行動を表現することが出来る。教示作業は、各階層の Categorizer の抽象化表現を定義することから始まり、定義された抽象表現に対する行動の写像関係の教示を最下層から段階的に行う。教示作業中に教示済みの TM によりロボットを行動させながら、その行動に対して修正を加える形で教示を行う。下層ほどセンサー観測値の抽象表現が単純なので簡単に教示が完了する。また、最下層以外では、すべての抽象表現に対して行動が割り当てられている必要がない。

7. シミュレーションによる検証

7.1 検証で扱った自律型海中ロボットモデル

以上の訓練システムの検証をシミュレーションにて行った。AUV の訓練を海中で行う事は困難なので、シミュレーションによる教示が有効である。AUV モデルを R-One Robot³⁾ とした。R-One Robot は、前進速度 3.6 ノット、垂直速度 0.5 ノット程度で、高度制御用にエレベータと垂直スラストを持つ。ピッチ角の制限が ± 20 度程度なので、それ以上傾斜のきつい海底斜面を回避するには垂直スラストを使用する必要がある。R-One Robot の行動アルゴリズムの認識パラメータを、定常機体座標速度、機体座標潮流値、レンジセンサー値、スラスト指示値、目標位置に対する相対位置とする。表 1 のようなレンジセンサー配置とした。行動モデルは図 3 に示す通りである。禁止行動算出モデルは表 2 のようにした。

表 1 R-One Robot のレンジセンサー配置 表 2 R-One Robot の禁止行動計算モデル

Sensor No.	Horizontal Direction[deg]	Vertical Direction[deg]
0	0	10
1	0	30
2	0	90

Sensor No.	危険値[m]	禁止する行動名
0	40.0	前進
1	30.0	機首下げ前進
2	20.0	鉛直下降

7.2 障害物回避用行動モデルの教示

縦運動による障害物回避行動の教示を行なった。モジュールの入力値を縦方向のレンジセンサー値、出力値を速度指令値とする。各階層の Categorizer が出力するレンジセンサー値の解釈を表 3 のように教示した。なお、教示作業は静水中で行なった。教示する行動の内容としては、図 4, 5 中にその大体の軌跡を示した通り「定高度でエレベータ制御により障害物を乗り越えていく」ことを教示者は意識しているが、急な坂ではエレベータのみでは対応出来ないため、この軌跡そのものが目標とはなり得ない。しかし、その点について教示が進むに連れて目標のイメージがはっきりしていき、海中ロボットが詳細に行動を学んで行く様を示す (図 4, 5)。

[最下層の教示後] 一通り山を越える事が出来るようにな

ったが、山登りを垂直上昇、山下りを機首下げ前進のみによって行うので、教示者のイメージからかけ離れた軌跡をとる。この段階で教示者は、斜面に沿ってエレベータを使って航行させる時間を増やそうと考える (図4)。

[中間層の教示後] 山登りの際に機首上げ前進することを覚えるなど、軌跡は多少理想のイメージに近づいた。(図4), ここで教示者は、ロボットが山を下降中に2つめの山に出会った場合の行動が気になり、その際の行動を確認した結果、2つめの山に出会った時点でロボットが上昇下降のループに陥っていたので (図5), 最上層の教示にてそれを扱うことにした。

[最上層の教示後] 山を登る際の機首上げを始めるタイミングを遅らせ、より斜面の近くで山に沿って上ることを覚え、2つめの山に出会った際に、谷の奥まで下らずに滑らかに下降から上昇へと転じさせるように教示した。また、下降時の階段状の軌道を修正した (図4, 5)。

学習の結果得られたアルゴリズムは、解析および修正可能であり、最下層の教示に不備があっても上層の教示でそれをカバーできる。

8. 結 論

潮流中における障害物回避による目的位置への直線航行を想定して自律型海中ロボットの訓練システムを構築し、シミュレーションによる検証を行なった。それにより、教

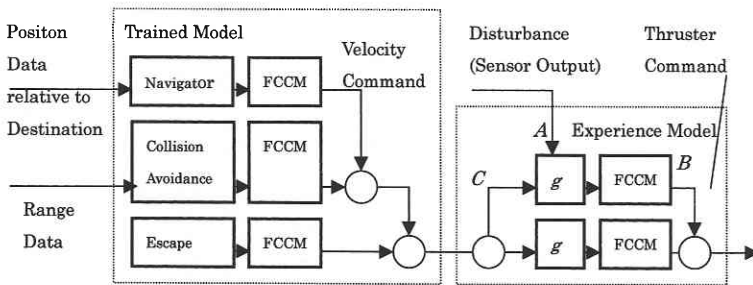


図3 検証システムの行動モデル

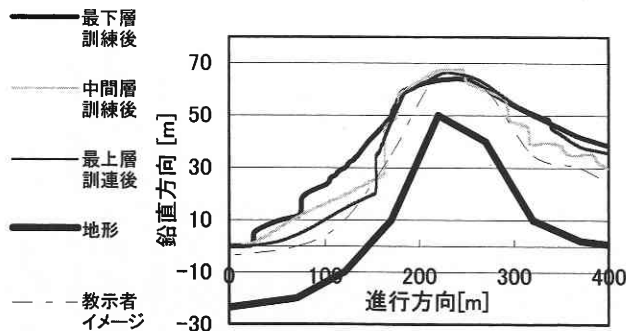


図4 障害物回避教示の結果 (山1つの場合)

示と自己学習を併用することで、教示の負荷を軽減することができることを示した。また、複雑な行動の教示を、信頼性を保ちながら行なうための手法を提案し、その有効性が示された。

(2000年6月10日受理)

参 考 文 献

- 1) 例えば Michael Sipser 著, 渡辺治, 太田和夫監訳「計算理論の基礎」共立出版 April, 2000.
- 2) Brooks, R. A.: "A robust layered control system for a mobile robot.", IEEE Journal of Robotics and Automation., RA-2, p. 14-23. April, 1986 .
- 3) 浦環, 小原敬史「航行型海中ロボット「アールワン・ロボット」の田辺市沖の自律潜航と海水観測」海洋調査技術第9巻 第1号 March, 1997.

表3 障害物回避行動モジュールのカテゴリライザの出力する抽象化表現

Sensor No.	Value = 0	Value = 1	Value = 2
0	0~40m	40~80m	80m~
1	0~30m	30~60m	60m~
2	0~20m	20~40m	40m~

Sensor No.	Bottom Category	Middle Category	Top Category
0 0 0	登り	崖登り	崖登り低
0 0 1			崖登り高
0 0 2			
0 1 0			崖登り低
0 1 1			崖登り高
0 1 2			
0 2 0			崖登り低
0 2 1			崖登り高
0 2 2			
1 0 0		登り坂	登り坂低
1 0 1			登り坂高
1 0 2			
1 1 0			登り坂低
1 1 1			登り坂高
1 1 2			
1 2 0	平地		
1 2 1		崖登り	崖登り低
1 2 2			崖登り高
2 0 0		山頂	山頂低
2 0 1			山頂高
2 0 2			
2 1 0		平地	平地低
2 1 1			平地高
2 1 2			
2 2 0	下り	下り坂	下り坂低
2 2 1			崖下り低
2 2 2			崖下り高

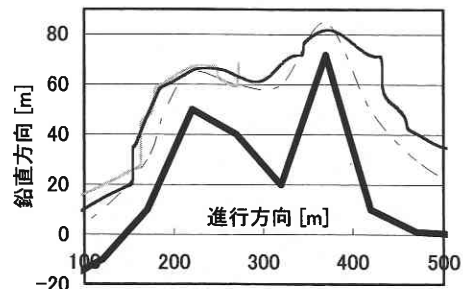


図5 障害物回避教示の結果 (山2つの場合)