

## 論文の内容の要旨

論文題目      Artificial Neural Network, Theory and  
Application to Dynamics Modeling and Advanced  
Control of Autonomous Underwater Vehicles

(和訳 自律型海中ロボットの運動特性の同定及び航  
行制御のためのニューラルネットワーク)

氏 名      サイヤディ, ハサン

本研究の目的は、自律型海中ロボットの動力学モデリングと、航行制御戦略の手法を提案することにある。各々の詳細な説明の前に強調すべきことは、本論文でとられている手法が、従来の運動モデルの同定や航行制御アルゴリズムの構築で使用されている方法とは異なる人工ニューラルネットワークによるものだけということである。本論文では、提案するモデリングと制御計画を自律型海中ロボットへ適用する。

自律型海中ロボットの航行制御において、時間変化のある多入力多出力非線型システムの動力学モデリングと正確なシステム同定は不可欠な要素である。剛体の運動方程式を導き、運動特性を評価する手法としては、ニュートンラグランジュ力学による手法を適用することが多い。航行制御への適用を想定した場合、ニュートンラグランジュの手法によって運動方程式を導くためには、正確な運動方程式の形を記述することに加え、プラントに与えられるすべての外力とモーメントの項を可能な限り明らかにする必要がある。海中ロボットのためにニュートン力学に基づく運動方程式がいくつか提案されているが、海水の流体力学的付加質量などの影響を同定することの難しさや、自律型海中ロボットのアクチュエータが持つ非線型特性などによって容易には適用することができない。

本論文では、そのような運動モデルの同定手法として、ニュートンラグランジュの手法に代わって、ニューラルネットワークによる汎用的な同定手法を提案し、これを Coupled Model Neural Network Identifier (CMNNI) と呼ぶことにする。入力、状態変化速度、力、モーメントであり、出力は状態変数とする。同定用のニューラルネットワークは巨大で複雑な物となり、これを一度に学習させるのは難しい。これを解決するために、6 自由度の運動モデルを表現する CMNNI を 6 つの部分 (Single Degree of Freedom Neural

network Identifiers : SDFNNI) に分ける事にする。

SDFNNI は 3 層のニューラルネットワークからなり、入力層は 2 つの入力用ニューロンと、6 つの中間層からの回帰結合用ニューロンからなる。中間層は、6 つのニューロンからなる。出力層は 1 つのニューロンからなり、その出力は 2 回積分される。各々の SDFNNI の学習は、評価関数の値が適当なレベルに収束するまで行われる。学習の済んだ 6 つの SDFNNI を、追加シナプスにより結合することによって 6 自由度の運動モデルである CMNNI を構成する。CMNNI において提案されるネットワークの構造は、どのような 6 自由度の剛体モデルにも適用可能である。

誤差伝播バックプロパゲーションによる教師付き学習の手法を、SDFNNI と CMNNI の両方の学習に適用した。教師付き学習では、教示信号としての入出力データセットを用意する必要があり、そのデータの収集のために、4 自由度の運動自由度を持つ自律型海中ロボットのツインバーガー 2 号を使用した。実験により各自由度の個々の運動特性と、各方向のカップリング特性のデータ収集をおこなった。運動データは、ツインバーガー 2 号に搭載のドップラ式流速計によって計測した。シミュレーションによる教師付き学習の結果が本論文において説明されている。

本研究において、自律型海中ロボットにおける航行制御の問題は、位置制御と経路計画の 2 つの観点から考えられている。はじめに、ニューラルネットワークを用いた正確なポジショニングについて説明し、その後で衝突回避と経路計画について述べる。位置制御モジュールは、海中ロボットを希望の位置に保持するための低レベルの制御モジュールとして扱う。経路計画モジュールは、障害物を回避するための経路を生成し、その結果を下位の位置制御モジュールに出力する。

自律型海中ロボットの位置制御は、海水の流体力学的な付加質量により、機体の力学特性が非線型なことから困難である。自律型海中ロボットのような多入力多出力の非線型なシステムを対象に、従来の制御理論による手法で制御パラメータやゲインを調整するのはほぼ不可能と思われる。自律型海中ロボットの航行制御に適用される制御システムには十分な適応性と、柔軟性が求められる。

本論文では、そのような位置制御の手法として、従来の制御理論による手法に変わるものとしてニューラルネットワークによる手法を提案し、Coupled Model Neural Network Controller(CMNNC)と呼ぶことにする。入力は、目標速度からの誤差、目標位置からの誤差、1 ステップ前の制御信号であり、出力は制御信号である。CMNNC のニューラルネットワークは巨大で複雑な物となり、これを一度に学習させるのは難しい。これを解決するために、6 自由度の運動モデルを表現する CMNNC を 6 つの部分 (Single Degree of Freedom Neural network : SDFNNC) に分ける事にする。

SDFNNC は 3 層のニューラルネットワークからなり、入力層は 2 つの入力用ニューロンと、1 つの出力層からの回帰結合用ニューロンからなる。中間層は、6 つのニューロンからなる。出力層は 1 つのニューロンからなる。各々の SDFNNI の学習は、評価関数の

値が適当なレベルに収束するまでおこなわれる。学習の済んだ 6 つの SDFNNI を、追加シナプスにより結合することによって 6 自由度の制御もジュールである CMNNC を構成する。CMNNC において提案されるネットワークの構造は、どのような多入力多出力制御システムにも適用可能である。

適応機能をニューラルネットワークで実現するにあたり、計算機の処理能力の問題から、制御用ネットワークの学習を、はじめはオフラインでおこなう。教師付きと教師なし学習の手法を SDFNNC の学習に適用し、教師なし学習で CMNNC を学習させる。誤差伝播バックプロパゲーション法によって、シナプスの重み値を更新し、評価関数の値を適当な値に収束させる。

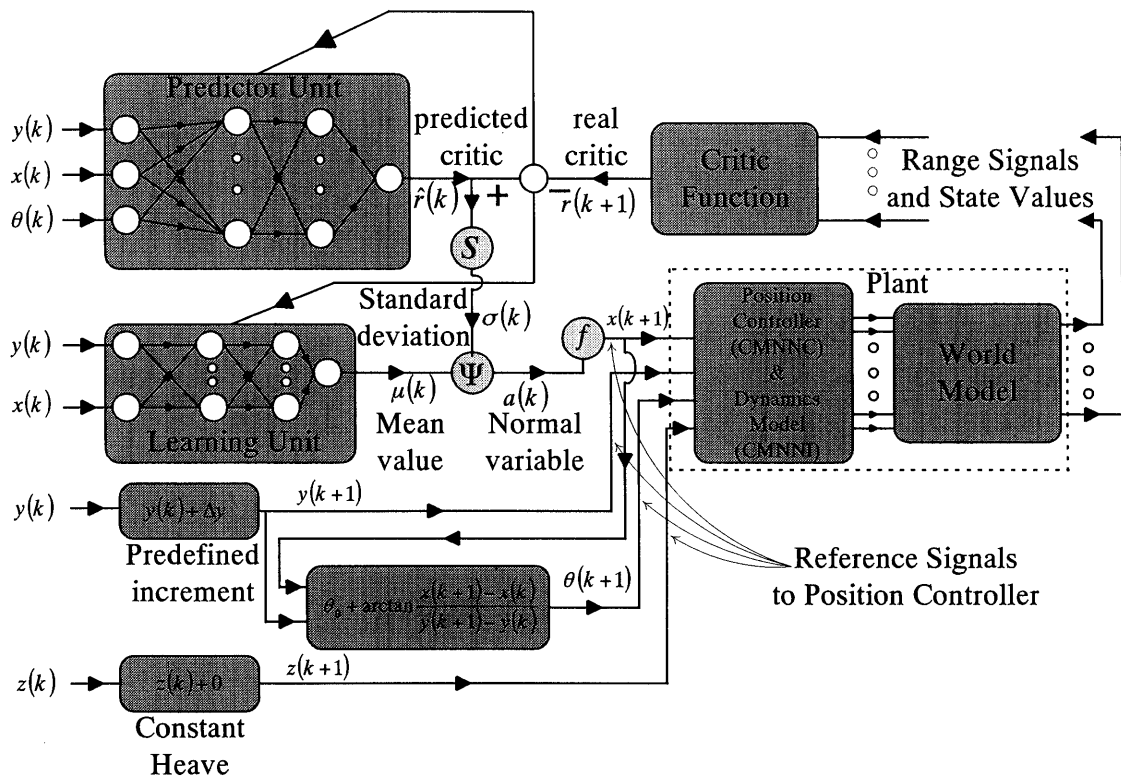
どのニューラルネットワークも、シナプスの重み値を更新するために、教示信号としてのデータセットを必要とするが、各 SDFNNC の学習の初期段階においては、別途用意された未熟なコントローラにより教示データを作成し、それによる教師付き学習をおこなった。

SDFNNC をより良いコントローラにチューニングするために教師なし学習をおこなった。CMNNC の学習用の未熟なコントローラの作成はほとんど不可能であるため、チューニング済みの各自由度の SDFNNC で代用し、それにより教師付き学習をおこなった後、教師なし学習で CMNNC のチューニングをおこなった。学習は、前述のツインバーガー 2 号の運動モデルと学習済みの SDFNNI と CMNNI によるシミュレーションによりオフラインでおこない結果を図示した。

衝突回避のための経路計画は本論文で考えるもうひとつの制御体系である。自律型海中ロボットのための障害物回避は非常に重要な要素であり、ミッションの途中でロボットが危険な環境物との衝突により障害を受けるのを防ぐために考慮されなければならないものである。自律型海中ロボットの運動計画は経路計画とトラジェクトリ計画に分解できる。経路計画衝突回避（モジュール）は、ある何らかの評価基準に関して最適化しながら、衝突の無い軌道を生成する。トラジェクトリ計画は自律型海中ロボットの運動を、計画された軌道に追従するように予定する。本論文では、前者の経路計画に焦点を絞る。

本論文で提案する自律型海中ロボットの経路計画衝突回避モジュールはニューラルネットワークによる航行制御の一例である。学習方法は強化学習に基づいており、連続した出力を生成する確率的な実数の関数を用いている。強化学習の基礎に基づいて、自律型海中ロボットが行動する環境から良い評価生み出す制御行動は報酬を受け、良い評価を生み出せない行動はペナルティを受ける。ここでは、所望する制御タスクに対して、確率的な強化学習を用いた経路計画衝突回避モジュールを適用した。システムは 2 つの部分からなる。ひとつは、入力に対して連続した値の出力を生成する学習ユニットで、もうひとつは、強化値を予測する役割の予測ユニットである。両方が同時に学習をおこなう。

ここで提案した確立的な強化学習手法は、自律型海中ロボット（例としてツインバーガー 2 号）の衝突回避行動のための平面経路計画に使われる。Fig.1 に示すように、このコ



**Fig. 1** Collision Avoidance module for AUV, Twin Burger 2, based on the Stochastic Reinforcement Learning method.

ントローラ・ネットワークは Coupled Model Neural Network Controller (CMNNC)と Coupled Model Neural Network Identifier (CMNNI)とにシリアルに結合されており、障害物回避のポジショニングのための目標値を生成する。適用した経路計画衝突回避の強化学習手法はただ 1 つのユニットを持ち、ロボットの平面運動に関する 3 つの異なる状態変数のうち、このモジュールにより 1 つのみが評価される。予測ユニットは、それぞれ 6 つのニューロンを持つ 2 つの隠された層を持つ。予測ユニットには平面運動における地球座標系の状態変数 3 つが入力される。学習ユニットもそれぞれ 6 つのニューロンを持つ 2 つの隠された層を持っている。学習ユニットには地球座標系における直交座標の 2 つの値が入力される。この 2 つの入力は学習ユニットが出力を生成するのに十分である。

地球座標系のロボット行動空間は平面であり、3 つのパラメータで表現される。行動空間は出発点から終点の方向にあらかじめ決められたある間隔で区切られ、強化学習モジュールはその直交方向の座標値を出力する。ロボットの回転運動は以上 2 つの値から計算される。経路計画衝突回避モジュールは現在ある位置から次の位置へのそれぞれの行程に、直交方向の適切な値を生成することが期待され、それゆえミッション全体を通した評価値は最小になる。

評価関数は評価値を出力し強化学習アルゴリズムの最適化プロセスに使われる。この関数は2つの異なる部分から構成される。ひとつは回避行動の部分であり、もうひとつはロボットを目標地点に向けつづける部分である。回避行動部への入力にはロボットから障害物への距離であり、目標点到達部への入力には横方向の位置である。評価関数は所望の行動を満足するように、これら2つに重みをかけた和で構成される。これら2つの適切な互譲が経路計画のアルゴリズムに障害物回避と目標点到達の両方を合理的に満足させる経路を生成させることができる。

本論文で提案するニューラルネットワークシステム以外のシステムと同様に、学習プロセスはシナプスの重みの初期化に関するある知識を必要とする。ある経路を経路計画モジュールのニューラルネットワークに初期経路として与えた。この経路が評価基準を満足するものではないことは明らかであるが、予測ユニットと学習ユニットのシナプスの重みをある初期値に調整するためだけにこれを用いる。学習の次の段階は最適化であり、確率的な強化学習手法により実現される。初期化された経路計画モジュールは評価値を用いて衝突回避をしながらの目標点到達のために最適化される。提案した確率的な強化学習手法を用いた経路計画衝突回避に関しては2つの事例研究をおこない、結果を図示した。

全体の結論として、本論文ではシステム同定と制御体系に、従来のモデリングや制御体系とは異なるアプローチを提案した。本論文において提案するモデリングと制御体系を自律型海中ロボットに適用した。

ニューラルネットワークによる自律型海中ロボットの運動モデリングを考えると、海中環境に由来する外力やモーメントを考慮する必要はない。これらのいくつかは非常に非線形であり測定することも容易ではない。加えて、自律型海中ロボットのスラストは容易にモデル化できない非線形性能を持っている。変数間の入出力関係を表現するデータを取得するための実験をすることが、モデルを構築するには十分である。ニューラルネットワークは入出力のマッピングを模倣するために、これらのデータセットを用いる。

同様に、ニューラルネットワークによる自律型海中ロボットの位置制御を考えると、コントローラのゲインを生成したり分析したりするのに古典制御の手法をとる必要はない。これらの手法を用いてコントローラを設計したり生成したり、コントローラの変数やゲインを調整するのは、このような非線形システムを扱う場合、とりわけ自律型海中ロボットのような多入力多出力システムの場合はほとんど不可能と思われる。ここではニューラルネットワークをこのような問題を扱うために用いた。学習を開始するために必要な初期状態のデータセットを生成するための非常に粗雑なコントローラが、学習手順の一番初めだけに必要である。その後は、いかなる追加データセットなしにコントローラは自動的に最適化される。

提案する運動モデリングや位置制御の手法は、自律型海中ロボットだけには限らないということをここに主張しておく。他の工学の分野における今日の制御プラントのほとんどは多入力多出力の類のものであり、これらのシステム同定や制御体系は提案する手法によ

り達成され得る。

自律型海中ロボットの衝突回避のための経路計画に関しては、自律型海中ロボット（ツインバーガー 2 号）の平面運動だけではあるが、提案した手法を適用した。本手法はこのロボットに限定するものではないと結論付けることができる。他のロボットシステムにも同様に適用が可能である。提案した最適化手法はニューラルネットワークを学習要素として用いた強化学習によるものである。このニューラルネットワークによる最適化手法もまた、他の制御タスクに適用することができる。制御問題に応じて、評価関数と本アルゴリズムで扱った環境の適当な変更が必要となる。