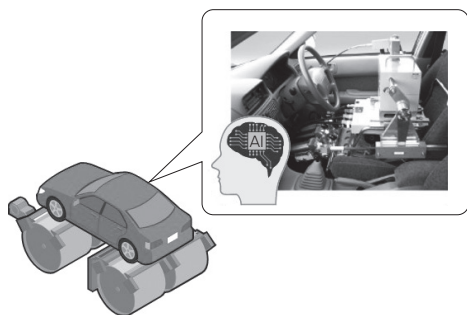


ドライブロボットへの人工知能（AI） 技術適用

吉田健人 Kento Yoshida
深井寛修 Hironobu Fukai
庭川 誠 Makoto Niwakawa
菅家正康 Masayasu Kanke

キーワード ドライブロボット制御、人工知能技術、深層学習、深層強化学習

概要



AI技術による制御
・指令車速への追従性向上
・人らしいペダル操作の実現

AI技術適用の効果

当社は、自動車試験の自動運転装置としてドライブロボットを開発している。ドライブロボットの制御性能への要求が高まってきており、指令車速への追従性の向上や人らしいペダル操作の実現を目的に、新たに人工知能（AI）技術による制御方法の導入を進めている。

教師あり学習によるアクセルペダルのフィードフォワード操作量の予測では、走行実績データの学習で過渡状態を含む車両特性を獲得することで、従来制御方法を上回る指令車速への追従性を獲得することができた。また、強化学習による制御方法の刷新では、指令車速への追従性の獲得のみならず、予測的なペダル選択を実現できた。

1 まえがき

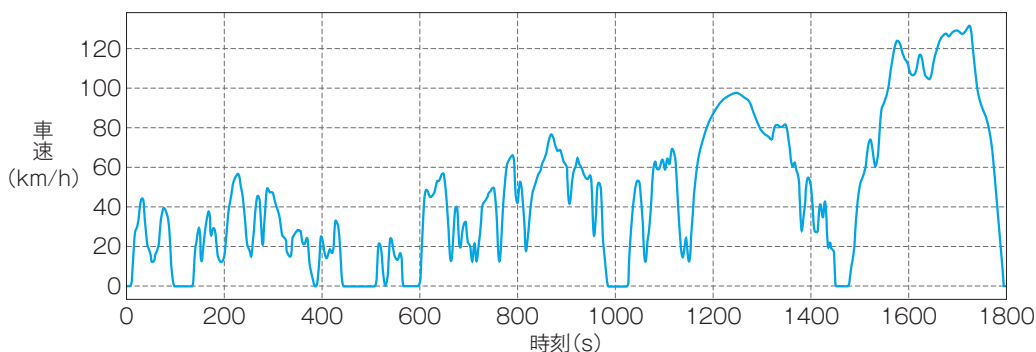
当社は、自動車の試験システムとして、路上での走行を台上で再現するシャシダイナモメータシステムを製造・販売している。ドライブロボットは、シャシダイナモメータ上での完成車両のアクセルペダル・ブレーキペダル・クラッチ・ミッションなどを操作する自動運転装置である。第1図にドライブロボットを示す。ロボットによる運転は、人のドライバーによる運転と比較して長時間の耐久運転や走行の再現性に優れている。自動車の性能評価のために様々な試験があるが、排ガス・燃費試験では、規格で定められた時間-車速パターン（以下、モード）に追従して走行した際の車両の性能を計測する。第2図にWLTC^(注1)モードを示す。試験に使用するモードは国・地域ごとに規格として定められるが、WLTCモードは国際的な基準として近年導入



第1図 ドライブロボット

完成車両の運転席に搭載され、アクセルペダルやブレーキペダルなどを操作する。

が進められている。ドライブロボットの基本的な機能要件としてもモードへの追従性が求められるが、近年追従性の更なる向上への要求が高まっている。



第2図 WLTCモード

横軸が時刻、縦軸が車速を表し、試験時にはこのような車速パターンに従った走行が求められる。

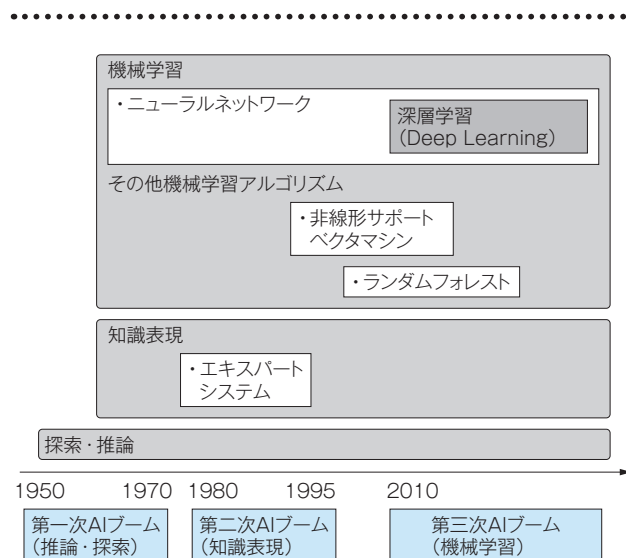
本稿では、プロドライバー並みの走行追従性や、より人らしい運転の実現を目的とし、人工知能（AI）技術をドライブロボットの制御に導入した事例を紹介する。

2 AI技術

2.1 AI技術の背景

AIとは、計算機によって人の知的行動を代替させる技術で、その概念は1950年代に初めて登場した。これまでに何度かブームと冬の時代を繰り返し、近年のブームは第三次AIブームと呼ばれている。第3図にAI技術体系を示す。過去のAIブームでは、推論・探索や知識表現などの手法が主流であったが、第三次AIブームは特に深層学習というニューラルネットワークによる機械学習技術にけん引されている。ニューラルネットワークは、過去のAIブームの時代にも存在していたが、時代とともに計算機の性能が発達したことなどで、過去には実現できなかった高い学習能力が認められるようになり、再注目されている。深層学習とは、大規模なニューラルネットワークと、それに上手く学習させるための一連のテクニックの総称である。

機械学習とは、データを利用してデータの背後にある何らかのルールを導く技術である。機械学習は、学習方法やデータの扱い方の観点から、大きく「教師あり学習」、「教師なし学習」、「強化学習」に大別される。教師あり学習とは、入力データと、それに対となる出力データを与え、入出力データの間に



第3図 AI技術体系

時代ごとに中心的なAI技術は移り変わっているが、現在に通じる技術の礎は古い時代にも存在している。

潜むルールを獲得し形式化（モデル化）する学習方法である。教師なし学習とは、与えられたデータ群の分布から、何らかの特徴を見出す方法である。強化学習とは、行動系の中で目指すべき状態基準を実現する行動の仕方を獲得する学習方法である。

深層学習は、それ単体で教師あり学習で使用されるのみならず、表現能力の高い関数近似器として強化学習の枠組みに組み合わせた使用方法も提案されてきており、この深層強化学習はAI分野の中でも近年特に注目度が増している。

以下に、ドライブロボット制御に深層学習による教師あり学習及び強化学習の適用を検討した取り組みを紹介する。教師あり学習は、制御系の中の車両特性収録部分に適用するが、従来制御の手法では反

映が困難であった車両特性も表現するように学習することで、制御性能の向上を目指す。また、強化学習適用では制御方法を抜本的に置き換え、モード走行というタスクの中で最適な車両の操作方法を学習プロセスによって自動的に獲得し、制御性能を向上させることを目的としている。さらに、フィードバック（FB）制御に見られる偏差を小さくするための瞬時的な操作の縮減が期待できる。

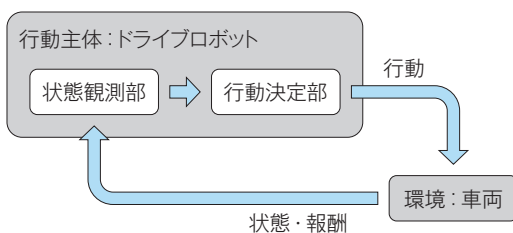
2.2 教師あり学習

教師あり学習では、入力データと教師データを用いて、機械学習モデルに入力データを与えた際の出力が教師データの値に近づくように、機械学習モデルを学習させる。学習後のモデルを用いることで、新たな入力データに対する予測出力を得ることができる。

学習の仕方や性能はアルゴリズムごとに異なるが、教師あり学習を実施するためには入力と教師のデータ組を用意する必要がある。ドライプロボット制御への適用では、車両の走行実績データを学習データとして使用し、指令車速とアクセルペダル操作の間の車両特性を獲得させる。

2.3 強化学習

第4図に強化学習の基本プロセスを示す。強化学習のシステムでは、大きく「環境」と「行動主体」が存在する。行動主体は環境の状態に基づいて何らかの行動を決定し、環境の中でその行動を実行する。行動主体の行動によって環境の状態が変化し、新たな状態を行動主体が観測して次の行動を決定する。行動主体は、状態の観測とともに「報酬」を受け取る。報酬とは、行動後の状態を評価したスカラー値である。強化学習では、一連の観測と行動を



第4図 強化学習の基本プロセス

経験を繰り返すことで、よい行動の仕方を獲得する。

繰り返した結果として、最終的により多くの報酬が得られる戦略的な行動の仕方を学習によって獲得する。ドライプロボットの制御の学習では、行動主体はドライプロボット、環境は車両となり、モードの指令車速との偏差が小さい検出車速であるほど大きい数値が得られる報酬を使用する。教師あり学習の場合には、学習後の制御特性は学習データの性質に依存するが、強化学習の場合には、例えば「人らしい運転」のような概念も学習基準として明示的に与えることができる。

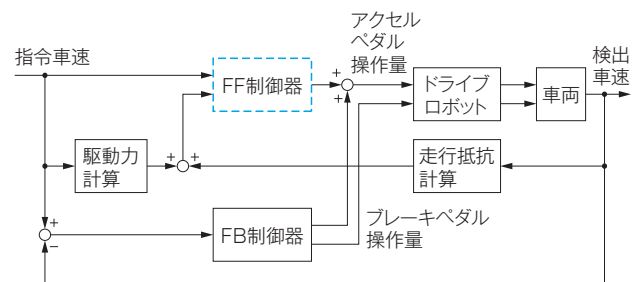
強化学習は、学習プロセスの中で環境の中での行動と観測を繰り返す必要がある。実物の行動主体を動かして学習することもできるが、シミュレータがあると学習させやすい。

3 AI技術の制御適用

3.1 教師あり学習の適用

ドライプロボット制御へのAI技術活用の一つとして、ニューラルネットワークで駆動力特性マップを代替した。

第5図にドライプロボットの従来制御の模式図を示す。ドライプロボットの従来制御では、モード走行のためのアクセルペダル操作量の決定に、駆動力特性マップというフィードフォワード（FF）系と、車速偏差のFB系を組み合わせた制御方法を用いている。駆動力特性マップは、モードの指令車速・指令車速を実現するために要求される車両の駆動力（要求駆動力）及びアクセルペダルの開度の関係を表したもので、試験実施時に定常的な車両特性を収録



第5図 ドライプロボットの従来制御模式図

アクセルペダルの制御はFF系（駆動力特性マップ）とFB系の組み合わせとなっている。

することで作成される。この駆動力特性マップに指令車速と要求駆動力を与えることで、アクセルペダルの操作量が得られる。駆動力特性マップの作成では、過渡的な車両特性は反映されていないため、実際の走行中に必要となるアクセルペダル操作量と誤差が生じる。そのためFB系によって、車速偏差を補償するアクセスペダル操作の補償量を出力する。

本取り組みでは、制御系の一部である駆動力特性マップをニューラルネットワークで置き換えた。

第5図中に点線で示したFF系をニューラルネットワークに置き換える。走行中の過渡的な状態も含めた車両特性を学習させることで、アクセルペダルの操作量予測の高精度化によるモード走行の追従性向上が期待できる。

ニューラルネットワークによる深層学習制御の評価のために、シミュレーションを実施した。ニューラルネットワークは、駆動力特性マップによる走行実績データを学習データとして使用し、検出車速・要求駆動力・エンジン回転数・将来一定時間の指令車速を入力して、必要なアクセルペダル操作量を出力するように学習した。なお、上位のアクセル/ブレーキペダル操作の切り替えや、ブレーキペダルの操作は従来制御と同じ方法を使用した。

第1表に各制御方法によるWLTCモードの走行結果を示す。モード走行試験では、ある時刻の指令車速に対して $\pm 1.0s$ 、 $\pm 2.0km/h$ の許容誤差範囲が定められているが、ファール時間とはその許容誤差範囲からの総逸脱時間を表している。従来制御でもファール時間は0sであったが、深層学習制御を適用することで、平均車速誤差・最大車速誤差ともに従来制御よりも小さくなった。

第6図にWLTCモードの一部区間の走行結果を示す。従来制御で車速偏差の大きかった箇所の追従に改善が見られた。また、アクセルペダル操作のFB系への依存が全体的に小さくなった。以上のように、深層学習制御によってFF系の予測精度が向上し、指令車速への追従性が向上したことが分かった。

一方、**第1表**でペダル操作回数は従来制御と比較して増加した。ペダル操作回数が多いことは、アクセルペダルとブレーキペダルの踏み替えが多い傾

第1表 各制御方法によるWLTCモードの走行結果

WLTCモードの1800s間の走行結果を示す。指令車速への追従性に関しては、ファールがなく、平均及び最大車速誤差は小さいほどよい。ペダル操作回数は、人らしい運転の指標の一つとして示している。

	ファール時間 (s)	平均車速誤差 (km/h)	最大車速誤差 (km/h)	アクセルペダル操作回数 (回)	ブレーキペダル操作回数 (回)
従来制御	0.00	0.44	4.27	38	41
深層学習制御 FF+FB	0.00	0.28	2.48	42	43
深層強化学習制御 FF	0.00	0.37	2.56	51	52
深層強化学習制御 FF+FB	0.00	0.30	1.72	48	49

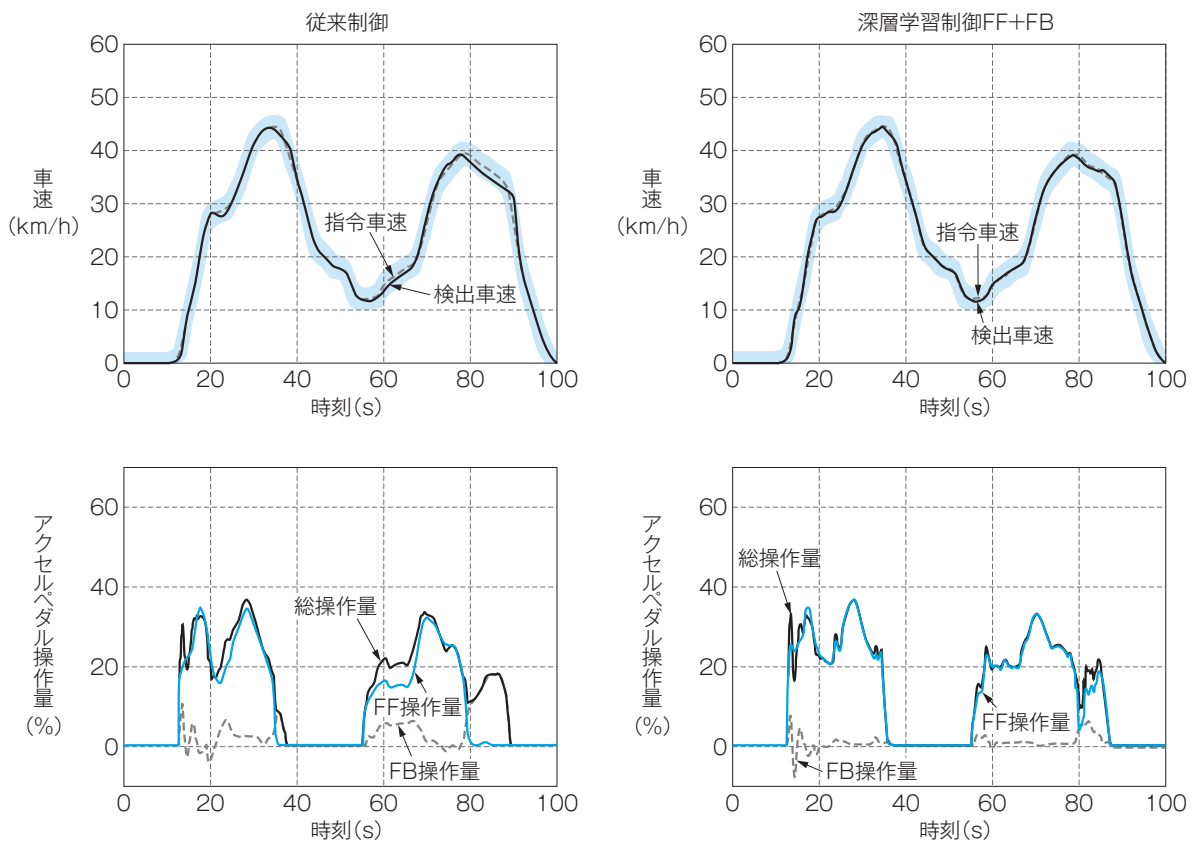
向を示す。追従性とペダルの踏み替え回数はトレードオフとなることがあるが、必要以上なペダル操作回数の増加は人らしい運転からの乖離につながる。また、**第6図**のアクセルペダル操作量の波形を見ると、深層学習制御では、ペダル操作の変化が大きい。急激なペダル操作の変化も、人らしい運転からの乖離となり得る。教師あり学習の場合、学習によって獲得できる制御性能は学習データの性質に依存しやすいため、人らしい運転の実現には人らしい走行実績データによる学習が必要と考えられる。

3.2 強化学習の適用

ドライブロボット制御へのAI技術適用の取り組みとして、深層強化学習の活用を検討している。本取り組みでは、モード走行というタスクの中で、アクセル及びブレーキペダルの最適な選択・操作方法の学習を試みる。

教師あり深層学習制御の適用では、従来制御方法を基本として部分的に機械学習に置き換えていたが、本取り組みでは制御方法を抜本的に変更することとなる。従来制御方法では、車両ごとのパラメータ調整や制御設計など、技術者による手動作業が必要であるが、強化学習システムでは、共通の学習プロセスの中で車両ごとの特性を吸収し、従来制御では設計しきれなかった制御ルールも暗に獲得できる。

深層強化学習制御の評価のために、シミュレー



第6図 WLTCモード一部区間の走行結果

上段：車速，下段：アクセルペダル操作量，左：従来制御，右：深層学習制御を示す。車速波形の塗りつぶしは，許容誤差範囲を表す。

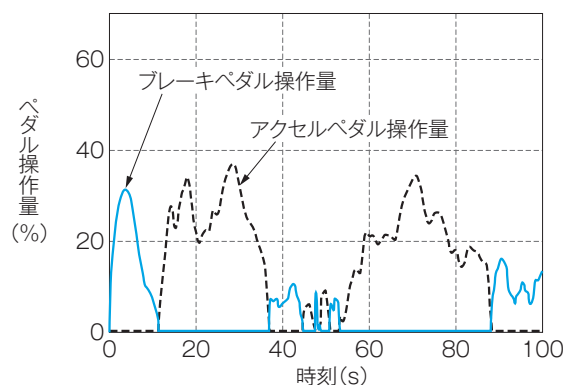
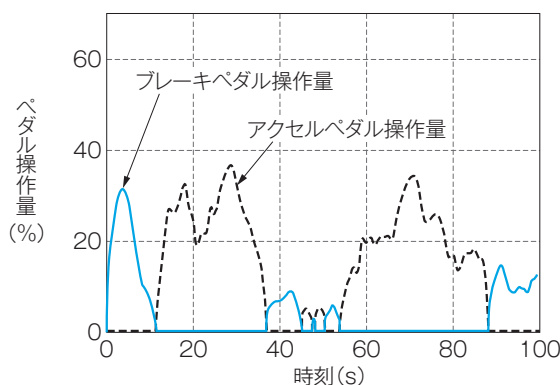
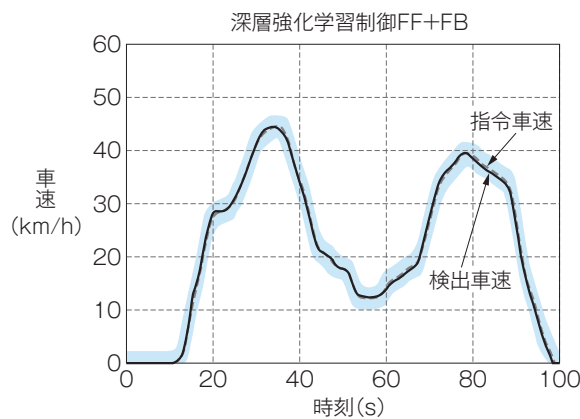
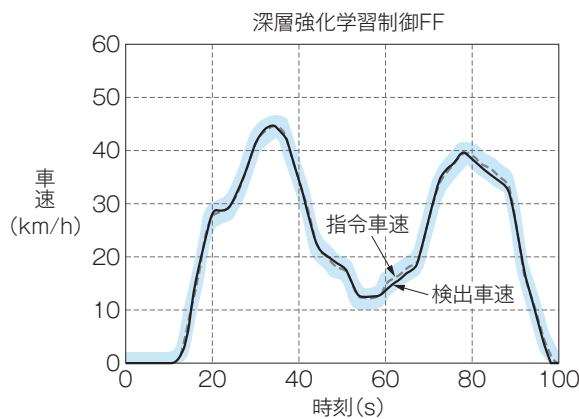
シオン実験を実施した。深層強化学習アルゴリズムでは、第4図の状態観測部から行動決定部の処理がニューラルネットワークで行われる。観測として、過去一定時間のアクセルペダル検出量・ブレーキペダル検出量・エンジン回転数・検出車速・指令車速・将来一定時間の指令車速をニューラルネットワークへ入力し、アクセルペダル操作量とブレーキペダル操作量を出力するように学習した。深層強化学習の状態評価基準である「報酬」は、以下の基準で定義した。

- (1) 指令車速への追従性が高いほどよい
 - (2) ペダル操作が滑らかであるほどよい
 - (3) ペダルの踏み替えがないほどよい
- 特に(1)の比重を大きくした。

第1表に示すように、FB系の有無にかかわらず、深層強化学習制御でもファール時間0sでモード走行できた。また、平均車速誤差・最大車速誤差ともに従来制御よりも小さく、指令車速への高い追従性能が得られた。特にFB系を組み合わせると、教師あり

学習適用時の走行結果よりも高い追従性能を示した。第7図に深層強化学習制御によるWLTCモードの一部区間の走行結果を示す。10s付近の加速時には、第6図に示す従来制御ではFB系が過敏にはたらい急激なアクセルペダル操作となり、オーバシユートが生じているが、深層強化学習制御では滑らかにアクセルペダルを踏み込みながら指令車速に追従している。これらのことから、深層強化学習制御では、ペダルの踏み込み量の推定にはまだ改良の余地があるが、操作すべきペダルを適切に予測できていることから、制御系への活用の有効性が示唆された。

一方で、教師あり深層学習制御の結果と比較しても、更にペダル操作回数が多い結果となった。第7図の48s付近にも0.5s間程のブレーキペダル操作があるなど、ごく短時間でペダルの踏み替えが何箇所かに見られた。また、FB系を組み合わせた場合には、ペダル操作の滑らかさも損なわれている。人らしい運転の実現のために、報酬設計や学習方法の見直しが今後の課題となる。



第7図 深層強化学習制御によるWLTCモード一部区間の走行結果

上段：車速，下段：ペダル操作量，左：深層強化学習制御FF，右：深層強化学習制御FF+FBを示す。車速波形の塗りつぶしは，許容誤差範囲を表す。

4 むすび

ドライブロボット制御へのAI技術の適用として、二種類の取り組みを紹介した。教師あり深層学習によるアクセルペダル操作量の予測では、走行実績データの学習で過渡状態も含む車両特性を獲得し、従来制御方法を上回る指令車速への追従性を獲得した。深層強化学習による制御では、指令車速への追従のみならず、指令車速に追従するための適切なペダル選択を実現した。

更なる追従性の向上とともに、ペダルの踏み替え回数の削減など、運転の人らしさの反映が今後の課題である。また、実試験時の学習方法を含むユーザビリティの検討にも取り組んでいく。

・本論文に記載されている会社名・製品名などは、それぞれの会社の商標又は登録商標である。

(注記)

注1. WLTC: Worldwide-harmonized Light vehicles Test Cycle (世界統一試験サイクル)

《執筆者紹介》



吉田 健人
Kento Yoshida
基盤技術研究所
人工知能技術の基礎研究開発に従事



深井 寛修
Hironobu Fukai
基盤技術研究所
人工知能技術の基礎研究開発及び画像処理技術応用製品の開発に従事



庭川 誠
Makoto Niwakawa
基盤技術研究所
人工知能技術の基礎研究開発及び画像処理技術応用製品の開発に従事



菅家 正康
Masayasu Kanke
動力計測システム工場
動力計応用製品の制御装置の開発に従事