

博士論文

ヒューマノイドにおける
搭乗型機器の運転操作と適応行動

平成30年6月8日提出
指導教員 稲葉 雅幸 教授

東京大学大学院 情報理工学系研究科
知能機械情報学専攻 博士課程
木村 航平

目次

第1章	序論	11
1.1	本研究の背景	13
1.2	本研究の目的	15
1.3	本論文の構成	17
第2章	オンライン探索指標に基づく搭乗型機器の運転操作と適応行動	21
2.1	はじめに	23
2.2	搭乗型機器の運転操作と適応行動における軌道修正とゲイン学習の必要性	23
2.2.1	運転操作における軌道修正の必要性	24
2.2.2	適応行動におけるゲイン学習の必要性	29
2.3	本研究における探索指標の定義と分類	34
2.3.1	探索指標の定義	34
2.3.2	探索指標の分類	34
2.4	オンライン探索指標の定式化	38
2.4.1	オンライン制御指標の定式化	40
2.4.2	オンライン学習指標の定式化	41
2.4.3	写像とブロック図	42
2.5	オンライン探索指標を利用した軌道修正とゲイン学習	44
2.5.1	オンライン探索指標に基づく軌道とゲインの一般化更新則	44
2.5.2	軌道修正とゲイン学習による探索目標へのアプローチ	46
2.6	搭乗型機器の運転操作と適応行動に対するオンライン探索指標のアプローチ	54
2.7	本章のまとめ	55
第3章	搭乗型機器の操作に対する軌道修正	59
3.1	はじめに	61
3.2	ヒューマノイドにおける閉ループ系の接触拘束	61
3.2.1	ヒューマノイド身体と接触対象の間に働く閉ループ系拘束	61
3.2.2	搭乗型機器の操作における閉ループ系の構造	64
3.3	複数のマニピュレータを協調させた搭乗型機器の操作行動戦略	66
3.3.1	三輪車の操作における課題	67
3.3.2	ヒューマノイドの双腕によるハンドルのステアリング操作	69

3.3.3	ヒューマノイドの双脚によるクランクのペダリング操作	69
3.3.4	ステアリング操作とペダリング操作における共通行動と特徴	72
3.3.5	探索指標の考えに基づいた操作行動戦略	74
3.4	探索指標を利用した能動機能と受動機能のマニピュレータ軌道における制御法	76
3.4.1	マニピュレータの能動機能と受動機能の定義	76
3.4.2	オフライン探索指標に基づく能動機能の制御則	78
3.4.3	オンライン探索指標に基づく受動機能の制御則	82
3.4.4	マニピュレータに対する探索指標の切替方策	86
3.5	ヒューマノイドの双腕・双脚による搭乗型機器の運転操作	88
3.5.1	実験準備	89
3.5.2	操作行動の実証実験	91
3.5.3	三輪車操作実験における能動・受動の機能考察	93
3.5.4	実践的な運転操作行動の実現	99
3.5.5	腱駆動ヒューマノイドによる運転操作	106
3.6	本章のまとめ	117
第4章	運転操作における搭乗姿勢と外界環境に対する軌道修正	119
4.1	はじめに	121
4.2	ヒューマノイドにおける搭乗姿勢の安定化と外界環境の認識	121
4.2.1	軌道修正が解く搭乗姿勢の安定化に対する課題	122
4.2.2	軌道修正が解く外界環境の認識に対する課題	123
4.3	搭乗姿勢の安定化に対する軌道修正法	125
4.3.1	運転操作時の着座姿勢のズレにおける問題点	126
4.3.2	割り込み処理型座り直し制御	128
4.3.3	実時間並行処理型姿勢制御	135
4.3.4	割り込み処理型座り直し制御と実時間並行処理型姿勢制御の違い	140
4.4	外界環境の認識に対する軌道修正法	141
4.4.1	外界環境における外乱要因と認識行動方策	141
4.4.2	大型障害環境に対する認識行動	143
4.4.3	小型障害環境に対する認識行動	156

4.5	本章のまとめ	159
第5章	搭乗する環境に適応したゲイン学習	163
5.1	はじめに	165
5.2	ヒューマノイドにおける足部トルクダンピング制御系	165
5.2.1	歩行系と搭乗環境適応系の違い	165
5.2.2	足部トルクダンピング制御系	168
5.3	ダンピングゲインのオンライン学習制御	170
5.3.1	SGD-based open-loop Learning Control (SLC)	171
5.3.2	SLCによるゲイン学習法	173
5.3.3	SLCに基づくオンライン学習制御系	176
5.4	平行二輪電動スクーターへの乗り込み動作	178
5.4.1	乗り込み動作におけるSLCの適用条件	178
5.4.2	SLCによるダンピングゲインのオンラインチューニングの実験結果	180
5.5	本章のまとめ	186
第6章	搭乗した環境の運動に適応したゲイン学習	189
6.1	はじめに	191
6.2	ヒューマノイドにおける調速行動制御系	191
6.2.1	搭乗した環境の運動	191
6.2.2	調速行動制御系	199
6.3	PIDゲインのオンライン学習制御	206
6.3.1	Mini-batch-based closed-loop Learning Control (MLC)	207
6.3.2	MLCによるゲイン学習法	208
6.3.3	前回値補正積分器	212
6.3.4	MLCおよび前回値補正積分器に基づくオンライン学習制御系	215
6.4	平行二輪電動スクーター上での調速行動	216
6.4.1	PIDゲインをオンラインで学習する調速行動実験	216
6.4.2	Iゲイン比に飽和制約を設けた方法と前回値補正積分器の比較	223
6.5	本章のまとめ	224

第7章	オンライン探索指標の並行処理と切替処理による行動実現	227
7.1	はじめに	229
7.2	オンライン探索指標の並行処理と切替処理	229
7.3	並行処理による行動実現	231
7.3.1	搭乗型機器の操作と実時間並行処理型姿勢制御に基づく並行処理実験	231
7.3.2	外界環境に対する認識と搭乗型機器への操作量逐次修正に基づく並行処理実験	234
7.3.3	目標 ZMP の軌道修正と調速行動のゲイン学習に基づく並行処理実験	239
7.4	切替処理による行動実現	249
7.4.1	搭乗型機器の運転中において割り込み処理型座り直し制御を実行する切替処理実験	249
7.4.2	搭乗型機器への乗り込み動作と搭乗後の調速行動をシーケンシャルに実行する切替処理実験	257
7.5	本章のまとめ	263
第8章	結論	265
8.1	本研究の総括	267
8.2	本研究の成果	268
8.2.1	本論文の手法により解かれる搭乗型機器の課題	268
8.2.2	本研究により得られる波及効果	271
8.3	本研究の結論	272
8.4	今後の課題と展望	273
	謝辞	275
	付録	279
A.1	搭乗型機器の構成	281
A.1.1	三輪車の仕様	281
A.1.2	平行二輪電動スクーターのモデル化	282
A.1.3	電動一輪車のモデル化	288
A.2	ヒューマノイドロボットの構成	291

A.2.1	等身大ヒューマノイドロボット“HRP2-JSK”	291
A.2.2	腱駆動ヒューマノイドロボット“腱志郎”	298
	発表文献	305
	参考文献	311

第1章

序論

1.1 本研究の背景

人間が行う作業の代行を目的として研究開発されてきたロボットは、これまでに数多くの課題を解決してきた。日本国内でロボットの研究開発およびロボット産業が活発化してきた1960年代、1970年代以降では、人間の身体部位のうち物理的な作業を中心的に担う手や腕などを機械的に模したマニピュレータやロボットアームが注目されてきた。研究開発においては、山下による人工の指の開発とハンドリングの実現 [1] をはじめ、井上による人工の手の計算機制御 [2]、内山による運動学の解析と制御システム構成、ビジュアルフィードバック制御処理について述べた人工の手の運動制御に関する研究 [3][4][5] 等が提案されている。また、産業分野での活用においても、川崎重工業株式会社が国内最初に導入した自動車組立ラインでスポット溶接作業を行うロボットアーム [6] や、株式会社安川電機が開発したティーチング・プレイバック型の全電気式産業用ロボット Motoman-L [7] など、工場等で人間が行っていた膨大な作業量を代行するロボットのマニピュレーション技術の恩恵を受けた。1980年代、1990年代に入ると計算機の進化とセンシング技術の発展に伴い、これまで机上で解析を行っていた複雑な制御理論などのロボット実機への適用がなされてきた。特に力学的なバランスの安定性を陽に考慮する必要のある二足歩行型ロボットの研究開発ではその恩恵を受け、下山が提案した竹馬形2足歩行ロボット BIPER-3 の動的歩行制御 [8] をはじめ、梶田らによる凹凸路面上の動的2足歩行制御 [9]、山口らが開発した多様な路面環境に適應できる2足歩行ロボットとその制御法 [10][11] が提案されてきている。2000年代以降は、これまでに培ってきたマニピュレーション技術と二足歩行型ロボットのバランス安定化制御を統合したプラットフォームであるヒューマノイドロボットの研究開発が盛んとなり、日本国内では ASIMO [12]、HRP-2 [13][14]、WABIAN-2 [15]、JAXON [16][17] など、国外では HUBO [18][19]、Atlas [20] などの等身大ヒューマノイドロボットが広く開発されている。同時に、ロボットの認識機能の向上や、深層学習 (Deep Learning) を中心とした機械学習手法に基づく人工知能技術の発達に伴い、知能を持ったヒューマノイドの研究開発も進みつつある [21][22][23][24][25][26][27]。

ヒューマノイド (Humanoid) はその名称の通り、人間 (Human) を模して開発されたロボットであり、人体模倣の身体的特徴を生かして人間が日常生活で利用している道具を普遍的に扱うことが可能な利点を有する。ヒューマノイドによる鋏、熊手などの複数道具の利用戦略獲得 [28] や、電動工具の操作 [29]、シャベルを用いた掘削作業 [30] 等の他、日常生活

空間に存在する多くの道具利用行動の実現がなされている。このようなヒューマノイドによる道具利用の研究は、半世紀ほどの歴史をもつマニピュレーション研究の延長線上に位置している。ある特定の作業タスクに特化した専用のロボット（マニピュレータやロボットアーム）を個々に開発していく従来の流れから、単一のロボット（ヒューマノイド）で複数のマニピュレーションタスクを行うために元来人間が利用してきた道具に適応しながら行動獲得していく流れが生まれた。

現在も引き続き研究されつつある道具利用によるマニピュレーションの拡張に比べて、道具利用によるロコモーションへの発展の実現は数少ない。ヒューマノイドロボット技術が発展してきた2000年以降では、HasunumaらやYokoiらによるフォークリフトやバックホーの遠隔操縦ヒューマノイドロボットを用いた運転実現 [31][32][33] が数少ない例であり、2013年-2015年に開催された災害対応を想定したロボット競技会 DARPA Robotics Challenge (DRC) [34] の課題タスクでは、ヒューマノイドによるユーティリティビークルの運転が最近になり実現された [35][36][37][38][39][40]。

例で挙げたヒューマノイドによる搭乗型の道具操作では、ヒューマノイド身体のうち腕部を用いたマニピュレーション（レバー操作やハンドル操舵）が主に作用するが、脚部の運動が支配的に作用する運転操作の行動実現例はほとんどない。ロボットが転倒しないよう固定した状態での単脚のみによるペダリング動作 [41] や、ヒューマノイドの足部をペダルに固定したビンディング環境での自転車のクランク回し操作 [42][43]、ロボットがバランスを崩さないよう人間の手で支えられた安定な状態での Pedal Racer 上でのペダリング [44] など安定的な環境を前提とした議論は存在しているが、脚部を活用した運転操作では、腕部のみによるマニピュレーションタスクと比べてヒューマノイドロボットの大きな課題であるバランスの安定化に関しても考慮に入れる必要がある。また、搭乗する道具に応じてヒューマノイドに実装した制御系統の調整、すなわち適応的な制御系のゲインチューニングも要求される。

搭乗型の道具を扱う際に特に重要となる課題は、ヒューマノイドの系が変わることへの対応である。通常のヒューマノイドのロコモーションは二足歩行系であるが、搭乗後は別のモビリティを獲得することとなる。外界の環境に対応してどのように動けばよいか、どのように道具を操作すればよいか、どれだけ制御系を調整すればよいかについて、ヒューマノイドと搭乗の対象物からなる一体化した系で議論することが重要となる。

本論文では搭乗型の道具を「搭乗型機器」と定義し、以上に述べた搭乗型機器に対する課題を解くためのヒューマノイドの運転操作と適応行動に関して論じるものとする。

1.2 本研究の目的

ヒューマノイドにおける搭乗型機器の運転操作と適応行動では、機器の操作・搭乗姿勢の安定化・外界環境の認識を目的とした (I) 軌道修正と、搭乗する環境への適応・搭乗した環境の運動への適応を目的とした制御系の (II) ゲイン学習が重要な柱となる。本研究では、(I) 軌道修正と (II) ゲイン学習の2つの方策をオンラインで実行することにより、ヒューマノイドの搭乗型機器に対する適応的な行動獲得の実現を目的とする。オンラインでの (I) 軌道修正と (II) ゲイン学習を実現するために、我々は共通の軸となる「オンライン探索指標」という考え方を提案する。オンライン探索指標の並行処理と切替処理を利用することで、搭乗型機器の運転操作と適応行動を最終的に実証する。(I) 軌道修正および (II) ゲイン学習の本質は以下の通りである。

(I) 軌道修正：

エラー（偏差）が0に近づくような目標の“軌道”を求めて“修正”する方策

(II) ゲイン学習：

評価関数（損失関数）が最小となるような制御系の“ゲイン”を“学習”してチューニングする方策

一般的な制御の考え方が、エラー（偏差）が0に近づくような入力を決定してプラントとなる制御対象に入力を与える方策であるのに対して、軌道修正は目標となる軌道を逐次修正する方策である。ここで“軌道 (trajectory)”とは、時間をパラメータとしてポーズ（位置と姿勢）を記述したものであり、ポーズ（位置と姿勢）の順序集合からなる“経路 (path)”と区別する [45]。ゲイン学習は制御系のゲインのセット・更新も含めたチューニングを含意する。

(I) 軌道修正と (II) ゲイン学習のそれぞれが目的とする搭乗型機器の運転操作と適応行動の実現課題を図 1.1 にまとめる。図 1.1 内の赤文字が課題を示し、本論文において対応する章も併記する。ヒューマノイドが搭乗型機器に対して適応的であるためには、事前に設計した計画軌道や予め設定した制御系ゲインのパラメータを、環境の変化に応じて逐次的に修正でき可変とするような行動戦略が求められる。

ヒューマノイドが作用するエンドエフェクタと、ハンドル等の搭乗型機器を操作するための対象物体は、事前に連結して設計されたものでなく、マニピュレーションで実行・再生される計画軌道と操作対象物体に拘束された運動との間の乖離を修正するためには、オンラ

インで軌道修正を行うことが不可欠である。さらに、搭乗型機器にはヒューマノイドが転倒しないよう常にバランスを監視しサポートする機能は当然備わっていないため、事前設計では把握できない振動などの外乱による転倒要因に対して、ヒューマノイド自身がオンラインで搭乗時の姿勢の安定化を図る軌道修正が必要となる。本研究では、ヒューマノイドの系および搭乗型機器の系の各々に対してのみのミクロな観点でみた軌道修正戦略にとどまらず、ヒューマノイドと搭乗型機器を併合した一体化した系についてもマクロな観点でみた軌道修正として外界環境を認識した行動戦略を議論する。

ヒューマノイドに適用された制御系のゲインの既定値の多くは平坦な床面等の環境を想定した設定値であるが、搭乗型機器に適応したゲインを探索し変更するチューニングが要求される。チューニング作業は人間がマニュアルで行うことが可能だが、ヒューマノイドが不安定な場合はチューニングの調整の度にロボットが転倒し、再度ロボットを搭乗型機器に乗せ直してリセット・リスタートを繰り返す煩雑な作業が課題となる。この問題に対して本研究では、自律的にゲインをチューニングするゲイン学習の方策を提案し、ヒューマノイドの系と搭乗型機器の系の間で働く搭乗環境に適応したゲイン学習と、2つの系を合わせた一体化した系の運動に適応したゲイン学習をそれぞれ議論する。

(I) 軌道修正と (II) ゲイン学習の2つの異なる概念を共通に扱うための一般化した考え方として、オンライン探索指標の概念を本論文の第2章で述べ、これを用いて軌道およびゲインの更新を図る。オンライン探索指標は並行的または断続的に適用でき、その並行処理と切替処理の性質を利用することで統合的な行動実現を目指す。本論文の第7章においてその実証の成果を示す。

本論文が目的とする「ヒューマノイドにおける搭乗型機器の運転操作と適応行動」に関する研究は、1.1節の背景で述べたマニピュレーション、バランス安定化制御、認識、学習をはじめとした知能ロボットが物理的に実世界へ作用し適応していくための行動獲得に必要な要素を横断的に含んでいる。これまで人間が周囲環境を変えてきたように、ロボットが外部の環境に触れて作用することで実世界の構成を変えていく研究は実現されてきているが、ロボットの自己身体の構成を変えるような道具利用形態の獲得の実現は発展途上にある。特にこれまではマニピュレーションを主体とした道具利用の獲得がなされており、ロコモーションへ展開した自己身体の拡張が本研究のテーマとなる。本論文により、ヒューマノイドロボットの移動形態への発展として波及効果が得られることを期待する。

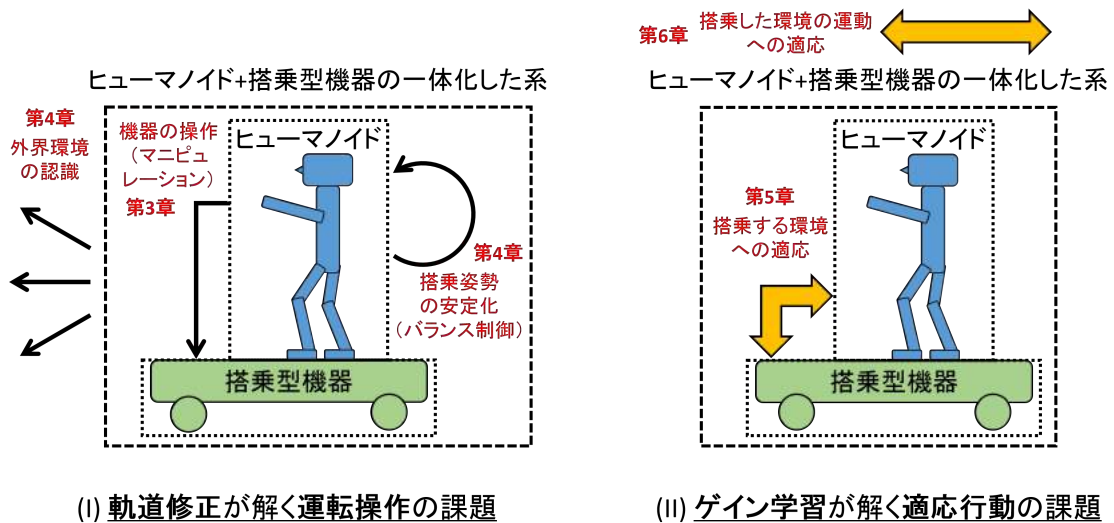


図 1.1: 軌道修正とゲイン学習が目的とするヒューマノイドにおける搭乗型機器の運転操作と適応行動の実現課題

軌道修正とゲイン学習がそれぞれ解くべき課題を赤字で示す。また、本論文での課題解決に対応する章を併記している。

1.3 本論文の構成

本論文を構成する章の全体の流れを図 1.2 に示す。本論文で提案するオンライン探索指標の考え方を主軸に、第 3 章から第 4 章では運転操作を目的とした軌道修正法を述べ、第 5 章から第 6 章では適応行動を目的としたゲイン学習法を述べる。以下で各章の概要について述べる。

第 1 章「序論」では、ヒューマノイドにおける搭乗型機器の運転操作（機器の操作・搭乗姿勢の安定化・外界環境の認識）を目的とした軌道修正と、適応行動（搭乗する環境への適応・搭乗した環境の運動への適応）を目的としたゲイン学習の重要性を述べ、それらの方策を一般化して扱うためのオンライン探索指標の考え方を基にした行動実現を目的とすることを述べた。

第 2 章「オンライン探索指標に基づく搭乗型機器の運転操作と適応行動」では、オンライン探索指標について定義を行った上で、オンライン探索指標に基づいた軌道修正とゲイン学習の戦略を述べ、搭乗型機器の運転操作と適応行動に対するアプローチを示す。

第 3 章「搭乗型機器の操作に対する軌道修正」では、搭乗型機器の操作におけるマニピュレーションに対しての軌道修正法を提案し、ヒューマノイドの双腕・双脚による搭乗型機器

の運転操作について論じる。

第4章「運転操作における搭乗姿勢と外界環境に対する軌道修正」では、運転操作時のヒューマノイドの搭乗姿勢の安定化を図る軌道修正戦略と外界環境の認識に基づく軌道修正戦略について述べる。

第5章「搭乗する環境に適応したゲイン学習」では、搭乗型機器の環境に適応するためのヒューマノイドの制御系の対象について述べ、対象となる制御系ゲインにおけるオンラインでのゲイン学習法を提案する。提案したゲイン学習法を搭乗型機器への乗り込み動作に適用し有効性を示す。

第6章「搭乗した環境の運動に適応したゲイン学習」では、搭乗後のヒューマノイドと搭乗型機器からなる一体化した系の運動について、速度に着目したヒューマノイドの制御系を新たに提案し、その制御系のゲインをオンラインで獲得するためのゲイン学習法について述べる。

第7章「オンライン探索指標の並行処理と切替処理による行動実現」では、主軸となるオンライン探索指標の実行処理について並行処理と切替処理の2つの処理方法を述べ、各処理に基づいた等身大ヒューマノイドロボットによる統合的な行動実現を実験により示す。

第8章「結論」では、本研究の総括ならびに得られた成果と結論を述べ、今後の展望について論じる。

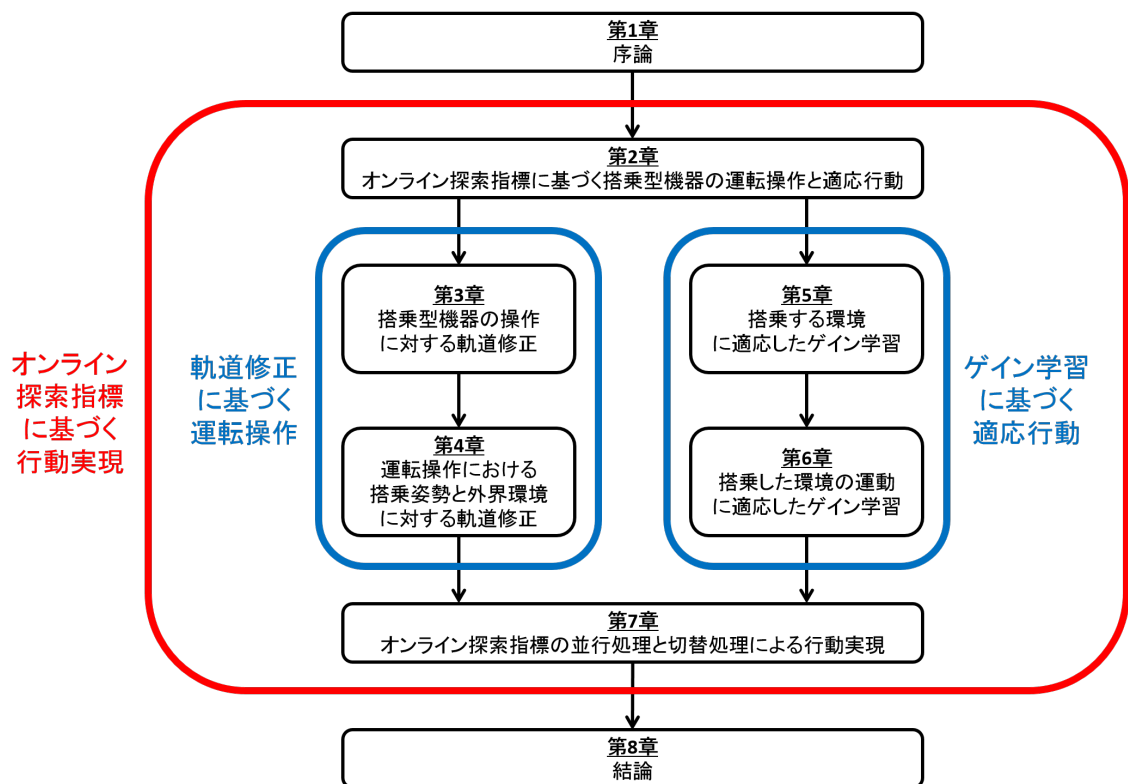


図 1.2: 本論文の章の構成

第2章

オンライン探索指標に基づく 搭乗型機器の運転操作と適応行動

2.1 はじめに

本章では，1.2節の本研究の目的で論じたヒューマノイドにおける搭乗型機器の運転操作と適応行動の課題を解くための軌道修正とゲイン学習の2つの異なる方策を共通に解くオンライン探索指標によるアプローチについて述べる．本章で提案するオンライン探索指標の作用と後続の本論文の章との対応を図2.1に示す．図2.1では運転操作と適応行動の2つの実現対象を全体集合で表し，その内の軌道修正およびゲイン学習により解かれる集合が各課題の部分集合を内包している．軌道修正が解く各課題の部分集合は

{ 機器の操作 (第3章), 搭乗姿勢の安定化 (第4章), 外界環境の認識 (第4章) }

からなり，ゲイン学習が解く各課題の部分集合は

{ 搭乗する環境への適応 (第5章), 搭乗した環境の運動への適応 (第6章) }

からなる．図2.1の最上部に位置するオンライン探索指標 (本章の内容) の集合は，軌道修正ならびにゲイン学習への写像

$$\begin{aligned} \text{オンライン探索指標} &\xrightarrow{f} \text{軌道修正} \\ \text{オンライン探索指標} &\xrightarrow{g} \text{ゲイン学習} \end{aligned}$$

を呈しており，単一のオンライン探索指標の集合から図1.1で示した5つの解くべき課題への対応が可能となる．

オンライン探索指標 \xrightarrow{f} { 機器の操作, 搭乗姿勢の安定化, 外界環境の認識 }

オンライン探索指標 \xrightarrow{g} { 搭乗する環境への適応, 搭乗した環境の運動への適応 }

また，図2.1の橙色の領域で示された各課題の共通部分が第7章の内容に相当し，オンライン探索指標の統合的な処理に基づき行動実現を達成させる．

2.2 搭乗型機器の運転操作と適応行動における軌道修正とゲイン学習の必要性

本研究の目的である搭乗型機器の運転操作と適応行動に対するヒューマノイドの軌道修正とゲイン学習に基づくアプローチの必要性について本節で述べる．本節の目的は，図2.1で示した運転操作と適応行動の全体集合における軌道修正とゲイン学習の集合および内包する各課題の部分集合の立ち位置を明らかにすることである．

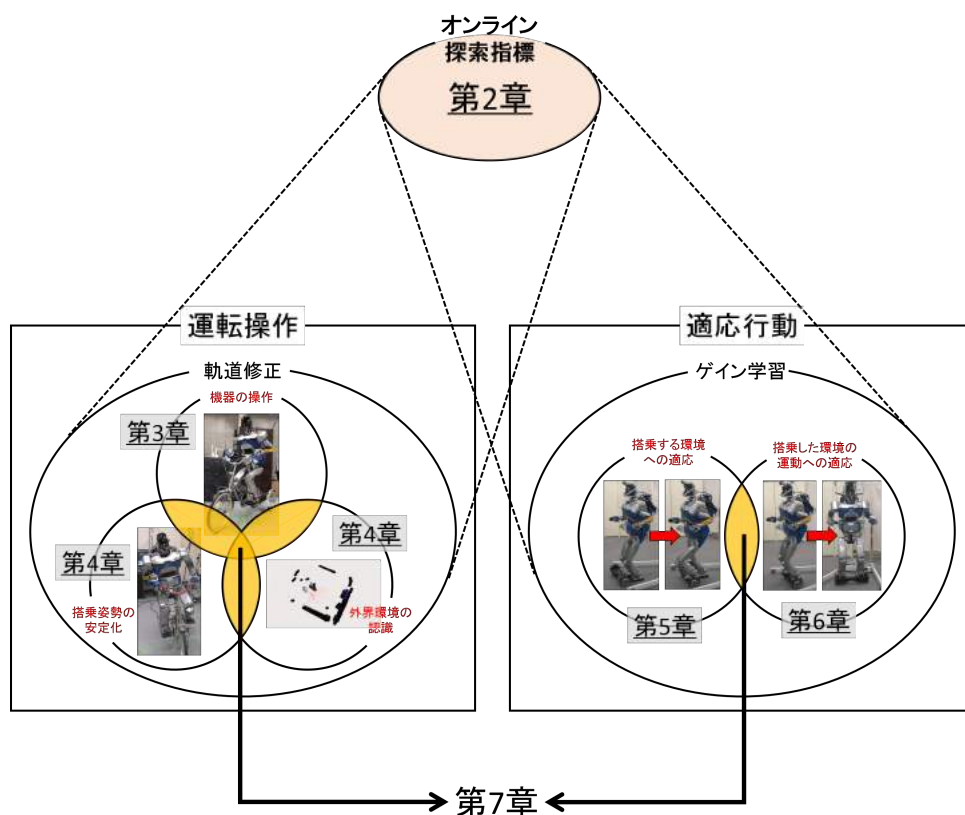


図 2.1: オンライン探索指標のアプローチによる運転操作の軌道修正と適応行動のゲイン学習への対応

共通となるオンライン探索指標の集合は、写像により運転操作（全体集合）の部分集合となる軌道修正と、適応行動（全体集合）の部分集合となるゲイン学習の 2 つの集合に写される。さらに、軌道修正の集合は { 機器の操作, 搭乗姿勢の安定化, 外界環境の認識 } の 3 つの課題の部分集合, ゲイン学習の集合は { 搭乗する環境への適応, 搭乗した環境の運動への適応 } の 2 つの課題の部分集合をそれぞれ内包している。オンライン探索指標を利用することで、上記 5 つの課題の共通部分（橙色の領域）の問題に対するアプローチも実現される。

2.2.1 運転操作における軌道修正の必要性

ヒューマノイドによる軌道修正の方策では、機器の操作・搭乗姿勢の安定化・外界環境の認識の 3 要素に対してのカバーが可能であることを先に述べた。これらの要素は具体的に以下の修正戦略ならびに図 2.2 に示す構成のように考えることができる。

- 機器の操作 ⇒ モデルと実環境の乖離の修正
- 搭乗姿勢の安定化 ⇒ 転倒因子に対する修正
- 外界環境の認識 ⇒ 障害を含んだ外界環境における操作軌道の修正

上記3つの修正戦略がそれぞれ解消すべき課題として、軌道を設計したモデルと実際の環境（実環境）との間で生じる乖離（エラー）、不安定なヒューマノイドロボットが転倒する要因（転倒因子）、ヒューマノイドと搭乗型機器が一体化した系の外部（外界環境）に存在する障害、が挙げられる。各々の課題に対して軌道修正が要求される必要性について、先行研究からの発展を踏まえて説明していく。

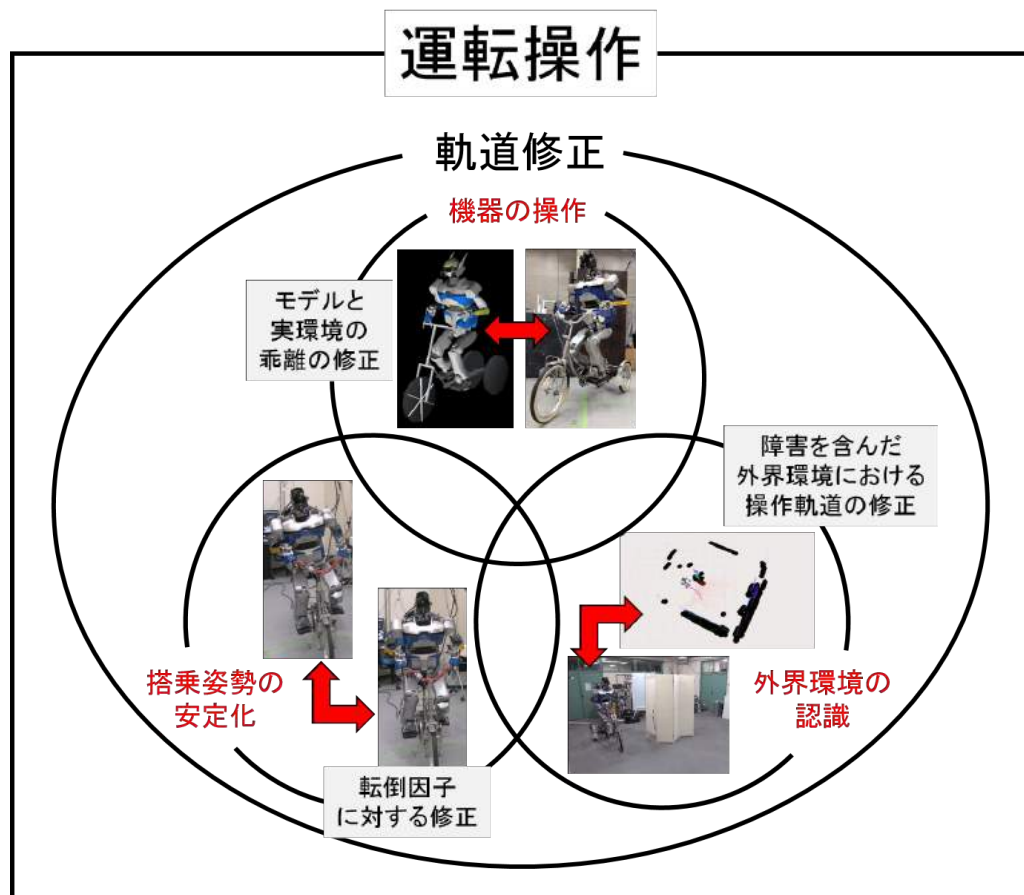


図 2.2: 運転操作において軌道修正が目的とする修正戦略

軌道修正が内包する機器の操作・搭乗姿勢の安定化・外界環境の認識の3つの部分集合は、モデルと実環境の乖離・転倒因子・外界環境における障害の各課題の解決を目的とする。

“機器の操作”に対する軌道修正の必要性

これまでにロボットは、シミュレーション等を利用してモデル化した結果と、実際の実機動作環境との間でズレが発生する問題に直面してきた。解決策のアプローチとして、一つ

は同定法やモデリングの精度を向上させる手法 [46][47] が提案されており、もう一つのアプローチはモデルと実環境との間でエラーが存在することを前提として、接触などの外乱に対してロバストとなるようロボットに搭載された力覚センサを用いた軌道修正法（インピーダンス制御をはじめとする力制御法 [48][49]，アドミッタンス制御手法 [50]）の他，カメラからの視覚情報に基づく逐次修正 [51] などが提案されている。

搭乗型機器の操作においても上記に挙げたような修正戦略が不可欠となる。図 2.3 の左側は，先行研究のヒューマノイドによるユーティリティービークルの運転行動 [37] におけるハンドルのステアリング操作を示している。操作行動では，事前に計画したヒューマノイドの手先（エンドエフェクタ）の軌道と，実際のユーティリティービークルのハンドルに拘束された運動（回転運動）との間にモデル化誤差や外乱が存在し，エラーの蓄積による目標操作とのズレ，ロボットの関節へのダメージが問題となる。先行研究 [37] では，この問題に対してインピーダンス制御を単腕に適用することで解決を試みている。

次に，ヒューマノイドによる三輪車の操作行動 [52][53][54] について考える。図 2.3 の右側は，ヒューマノイドの双腕による三輪車のハンドルのステアリング操作と双脚によるクラシクのペダリング操作を表す。ユーティリティービークルのステアリング操作との違いは，2つのマニピュレータを用いた協調マニピュレーションに拡張されている点である。単腕によるマニピュレーションと比較して，1つのマニピュレータと操作対象物体との間の乖離のみならず，双方のマニピュレータが互いに干渉し影響し合う問題に関しても考慮する必要がある。[37] で適用されているようなインピーダンス制御が，ロボットと環境・物体との接触の問題に大きな福音をもたらすのに対して，ロボットのマニピュレータ同士の軌道が干渉する問題に対してはマニピュレータの軌道修正が必要不可欠となる。十分に低い剛性のゲインを適用したインピーダンス制御はロボットの関節などへの負担を大きく軽減できるが，目標とする操作力と実際に発生する操作力との間のエラーを許容することに等しく，負荷は解消される一方で目標の操作行動が達成できない問題がトレードオフとして生じる。双腕による三輪車のステアリング操作の例では目標のハンドル角度まで操舵できない問題，双脚によるペダリング操作の例ではペダルを十分に踏み込めない問題が提起される。以上より，事前に計画した軌道から逐次的に修正を実行することが望ましいと考えられる。

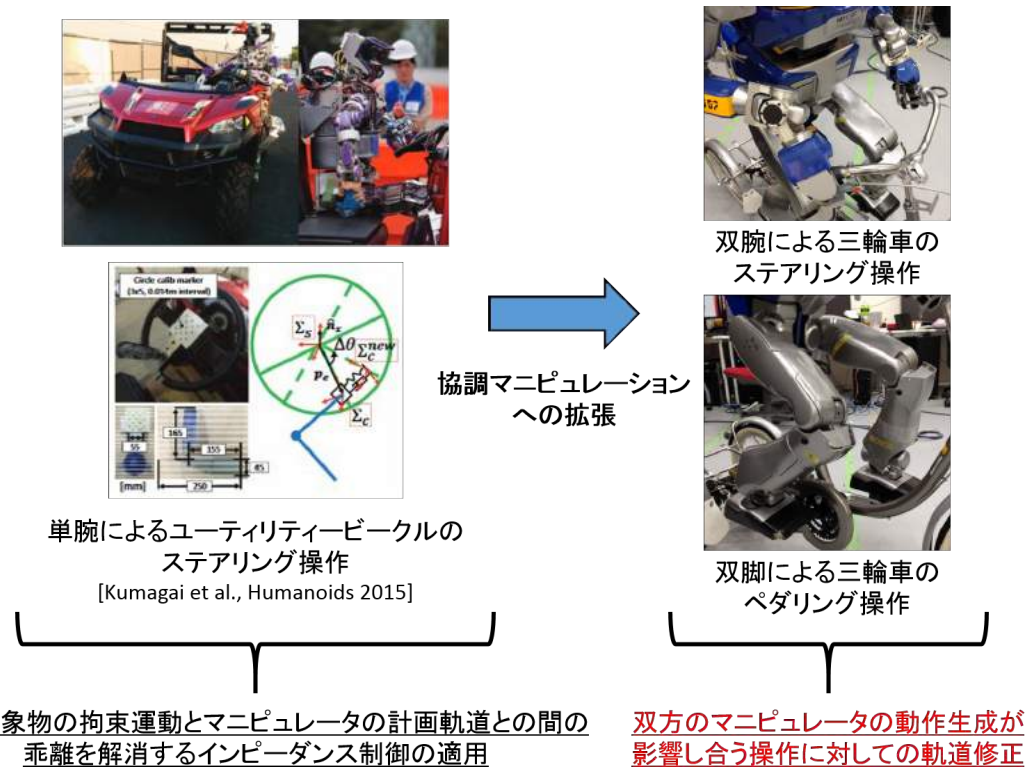


図 2.3: ヒューマノイドにおける搭乗型機器の操作に対する軌道修正

左側に示す先行研究 [37] の例では、インピーダンス制御を適用することで操作対象物（ハンドル）に拘束された運動（回転運動）と単腕マニピュレータの計画された軌道との間の乖離を解消している。右側に示す三輪車操作 [52][53][54] の例では、操作対象物（ハンドル・クランク）の拘束運動への対処に加えて、双方のマニピュレータ（双腕・双脚）の軌道が干渉し影響し合う問題に対して軌道修正を適用する必要がある。

“搭乗姿勢の安定化”に対する軌道修正の必要性

転倒を防止するバランスの安定化は、ヒューマノイドをはじめとする二足歩行ロボットの永劫的な課題である。二足歩行ロボットにおけるバランス制御法の提案ならびに実装はこれまで広く取り組まれている [55][56][57]。制御によるバランス安定化と異なるアプローチとして、ハードウェアの機構的变化により安定性の高い構造へと変形する行動アプローチも提案されている [19]。

ヒューマノイドが搭乗型機器を扱うことの利点の一つが状態の安定化の獲得である。[19]のように二足歩行形態から台車型の安定的な形態へと身体プロパティを変化することが可能である。図 2.4 の左側（バックホーの運転 [32]、車の運転 [40]）に示すように、ヒューマノイドが搭乗型機器に着座することで安定した姿勢でのロコモーションへと拡張がなされる。一

方で、図 2.4 の右側（三輪車の運転 [52][53][54]，平行二輪電動スクーター [58] の運転 [59]）に示す例においては，搭乗型機器上においてヒューマノイドは不安定な姿勢であり，従来の二足歩行ロボットのバランス制御法とは別のアプローチを考える必要がある．したがって，これらの行動例では搭乗型機器上でヒューマノイドの搭乗姿勢の安定化を図る目的として，姿勢や重心（Center of Mass: CoM）・ZMP（Zero Moment Point）[60] に対する軌道修正戦略が求められる．

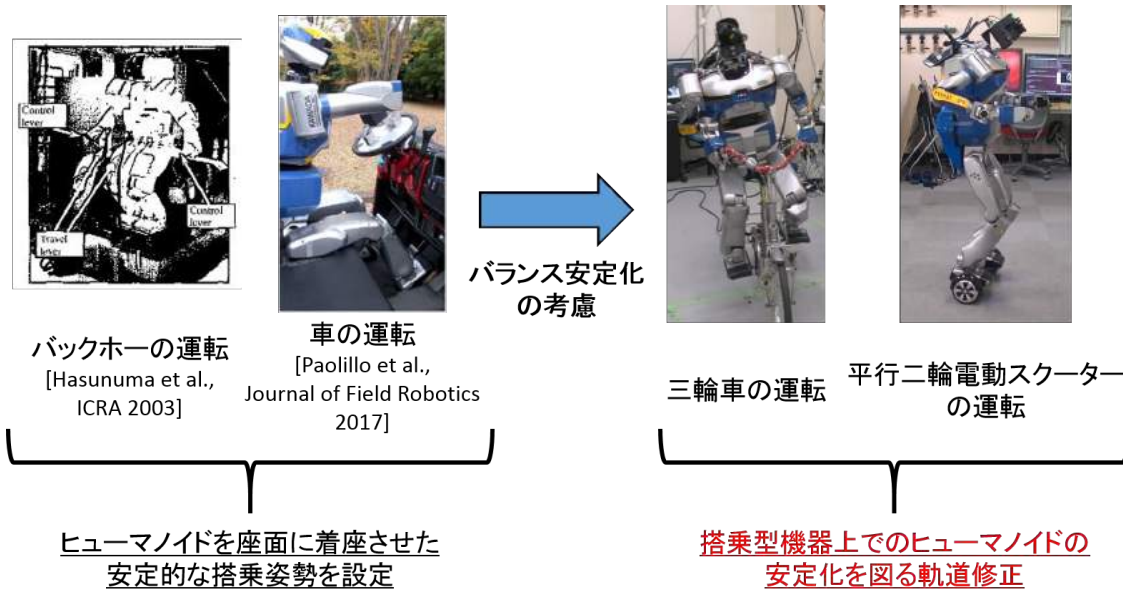


図 2.4: 搭乗姿勢のバランス安定化を目的とした軌道修正

左側 [32][40] の例では，ヒューマノイドを搭乗型機器の座面に着座させることで安定的な搭乗姿勢を確保している．右側 [52][53][54][59] の例においては，搭乗型機器に対してヒューマノイドが不安定となる状態を含んでおり，バランス安定化を図る目的の軌道修正が要求される．

“外界環境の認識” に対する軌道修正の必要性

ロボットに搭載されたステレオカメラによる視覚情報，LRF（Laser Range Finder）で測定される距離情報などを統合した SLAM（Simultaneous Localization and Mapping）[61][62][63][64][65] の技術や，触覚センサ，力センサを用いた接触の力覚情報に基づく未知物体・未知環境の手探り行動 [66][67][68] 等，ヒューマノイドの外界環境に対する認識法が広く提案されてきている．認識された環境の結果に基づいて，ヒューマノイドは行動の軌道を決定していくことになるが，搭乗型機器との併合によりヒューマノイドは新たな身体を獲得し，ロコモーション

ンの行動形態が変化することとなる。

一般的なヒューマノイドの外界環境認識行動は図 2.5 の左側（ヒューマノイドにおける障害物回避行動 [69]）に示すように二足歩行での移動形態（モビリティ）となる。そこから発展して、ヒューマノイドが搭乗型機器に搭乗することでモビリティの変化が起こり、この変化に対応した行動戦略が必要となる。図 2.5 の右側の行動例はヒューマノイドが三輪車を操作する場合の障害物回避行動 [53] を表し、外界環境における障害物と干渉しないような軌道修正が求められる。三輪車の操作行動の場合、どれだけハンドルを操舵すればよいか、どの程度漕ぎ進めればよいか、に関して搭乗するヒューマノイドが行動決定を行う必要がある。

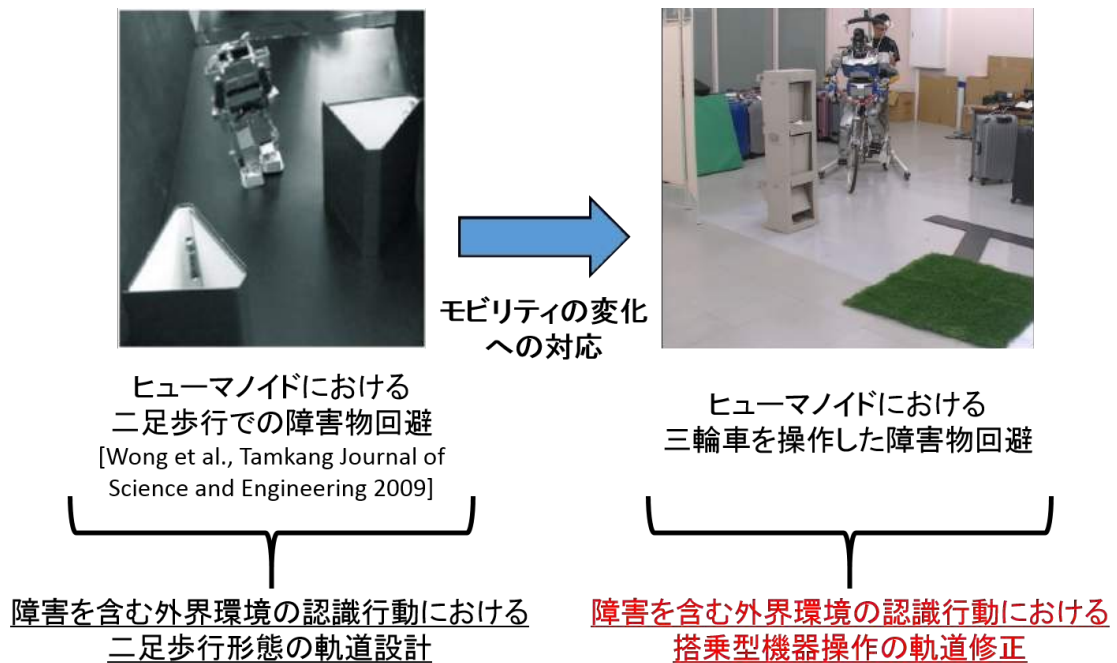


図 2.5: ヒューマノイドにおける外界環境の認識行動の軌道修正

左側 [69] は、ヒューマノイドの一般的な従来の二足歩行形態による外界環境認識行動（障害物回避）を表す。搭乗型機器への搭乗により、ヒューマノイドの移動形態（モビリティ）が変化（右側）し、これに対応した操舵や進行の行動決定における軌道修正が必要となる。

2.2.2 適応行動におけるゲイン学習の必要性

ヒューマノイドにおける適応行動では、搭乗する環境への適応・搭乗した環境の運動への適応の2要素に対するカバーが可能であることを先に述べた。これらの2つの要素は具体的に次の学習戦略ならびに図 2.6 に示す構成のように考えることができる。

- 搭乗する環境への適応 ⇒ 床面から搭乗型機器上へと環境変化を伴う際の適応的なゲイン学習
- 搭乗した環境の運動への適応 ⇒ 搭乗環境変化後の機器の運動に適応したゲイン学習

上記2つの学習戦略がそれぞれ解消すべき課題として、床面と搭乗型機器上の環境の違い、搭乗型機器の運動特性への対応、が挙げられる。各々の課題に対してゲイン学習が要求される必要性について、先行研究からの発展を踏まえて説明していく。

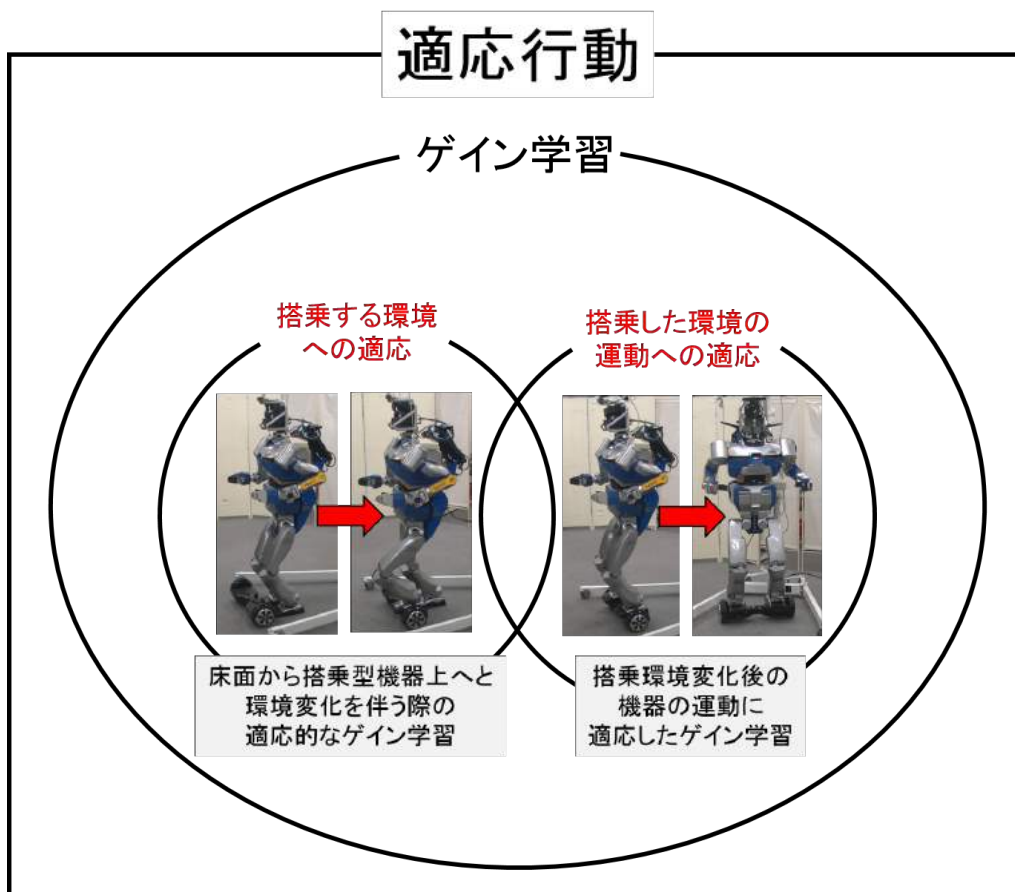


図 2.6: 適応行動においてゲイン学習が目的とする学習戦略

ゲイン学習が内包する搭乗する環境への適応・搭乗した環境の運動への適応の2つの部分集合は、床面から搭乗型機器上への環境変化・搭乗後の機器の運動への対応の各課題の解決を目的とする。

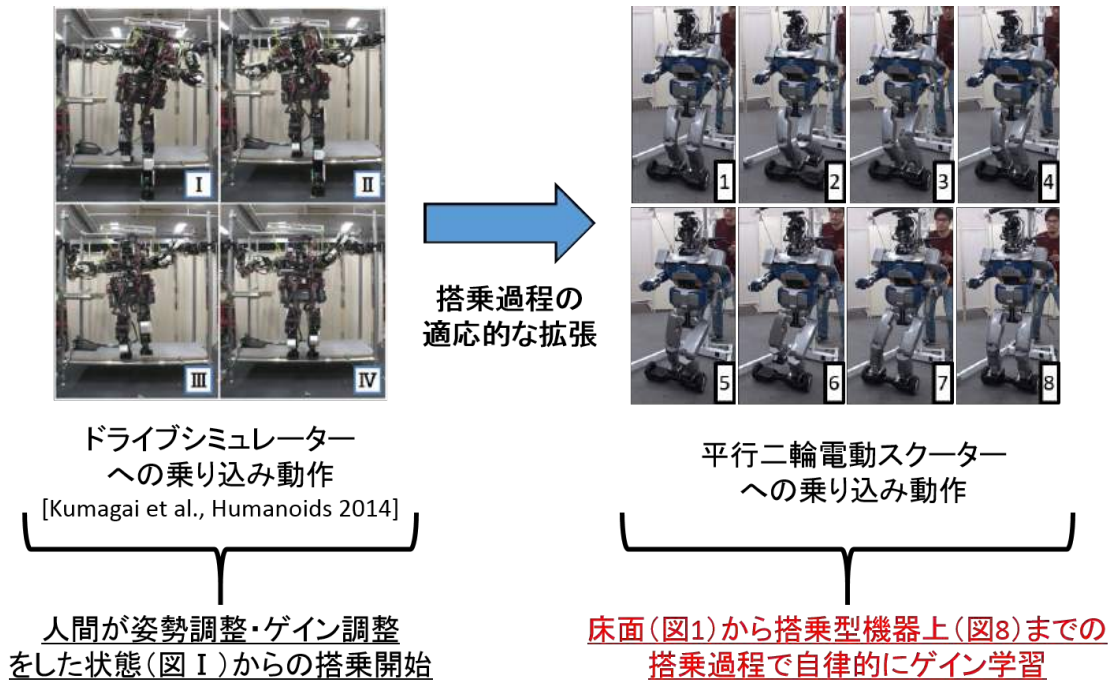


図 2.7: 搭乗する環境への適応を目的としたゲイン学習

左側に示すドライブシミュレーターへの乗り込み動作 [70] の例では、事前の調整に基づく部分的な搭乗過程において移乗行動を生成している。右側に示すような搭乗過程を拡張した床面から搭乗型機器上への乗り込み動作の例においては、搭乗する環境の変化に対して適応可能なゲイン学習が求められる。

“搭乗する環境への適応” に対するゲイン学習の必要性

ヒューマノイドが搭乗型機器へと移乗行動を実現するためには、乗り込み動作における姿勢列の生成に加えて、移乗前と移乗後の環境の違いによる制御系のゲインの変更が必要となる。制御系のゲイン調整は、各々の環境下におけるロボット実機の挙動を観察しながら人間が適宜チューニングを行うことが多く、シミュレーションではカバーできない点が課題となる。そのため、等身大ヒューマノイドロボットをはじめとする大型実機を用いて調整を行う場合、試行実験の繰り返しによる煩雑な作業（重量のあるロボットの実験環境への再設定など）が大きな負担となる。

図 2.7 の左側に示す大型の大出力ヒューマノイドロボット STARO [71] によるドライブシミュレーターへの乗り込み動作 [70] では、事前に片脚を機器上に乗せた姿勢を作成し、それに応じた制御系のゲインを与えた上で乗り込みを実行している。一方で、図 2.7 の右側に示す平行二輪電動スクーター [58] への乗り込み動作のように、平坦な床面で二足直立した

状態から機器上に移乗するまでの拡張された搭乗過程においては、フェーズ毎にゲインを適応的に調整する必要があり、即時的な乗り込み動作では人間による瞬時のチューニングが困難である。したがって、瞬時の乗り込み期間で自律的にゲインを学習させて更新する学習戦略が求められる。

“搭乗した環境の運動への適応”に対するゲイン学習の必要性

移動ロボットの開発においては、初期の設計段階でロボット下部の構成をどのようにすべきか（台車型ロボット [72]，倒立振り子型ロボット [73]，二脚型ロボット [74]，多脚型ロボット [75] 等）を決定する。その後の動作確認の段階では、一意に決定したロボット下半身の構成に対して十分な移動性が満たされるように、実装した制御系の調整を人間が行っていく。搭乗型機器に搭乗させたヒューマノイドのように設計・開発の段階では移動性の変容に関して考慮されていない場合は、搭乗状態にある一体化されたヒューマノイドに対して別個に調整を行う必要がある。

「搭乗する」という意義は、ロボットと搭乗型機器がハードウェアとして一体となったアタッチメント（図 2.8 左側の例 [76][77]）を形成するのではなく、互いの環境に接して触れている状態を表す。したがって、厳密にロボットと搭乗型機器の接合点が設計されている状態ではないため、搭乗したときの接触状態には座面とのズレや足の着地位置のズレなどのノイズ・外乱等の変動要素が含まれている。変動要素は搭乗状態にあるヒューマノイドの運動の特性に影響を及ぼすことから、変化に対処する制御系のゲイン調整が要求される。接合点が厳密に一意に決定している場合には人間による事前のチューニングで解決されうるが、搭乗した状態に応じて変化する場合（図 2.8 右側）は逐次的に人間がその都度調整を行うことは困難となる。上記のような搭乗した環境の運動に対しても適応的であるためには、ゲイン学習が必要不可欠となる。

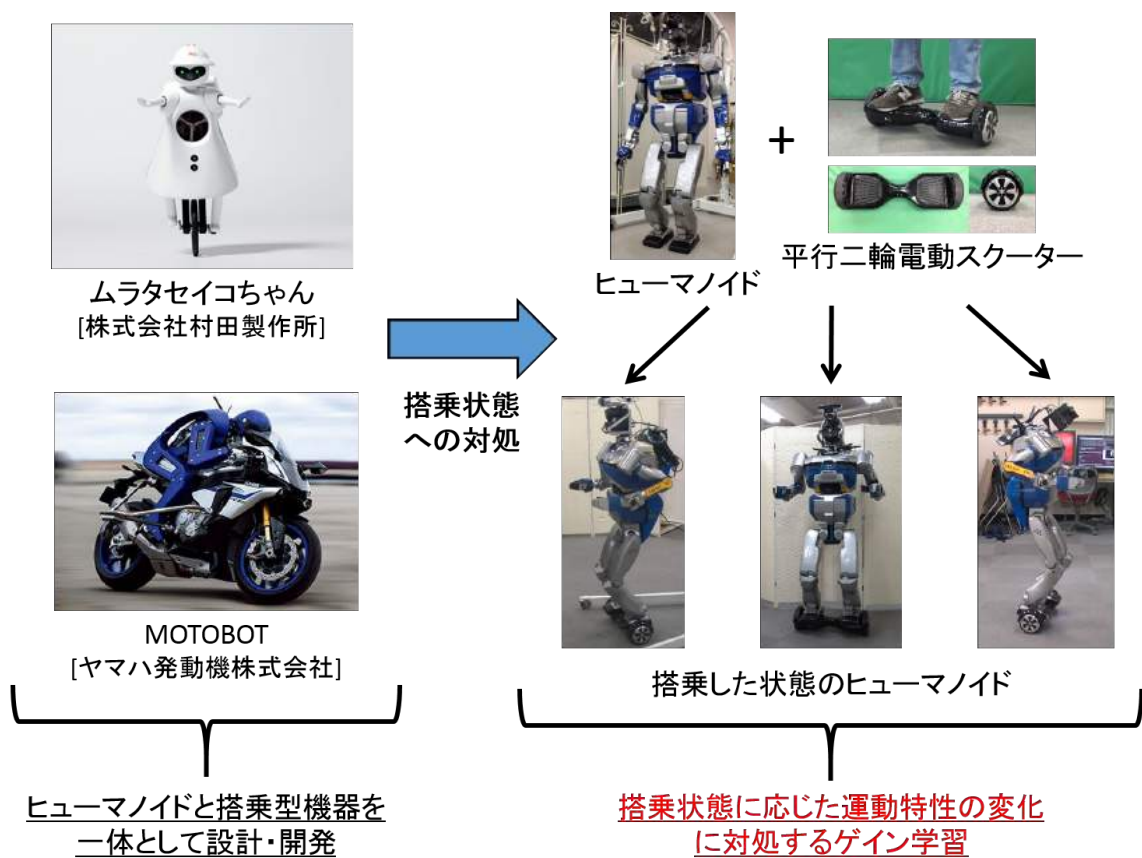


図 2.8: 搭乗した環境の運動への適応を目的としたゲイン学習

左側に示すロボットの例 [76][77] では、ヒューマノイドと搭乗型機器（一輪車やモーターサイクル）を一体として組み付け可能なように設計・開発を行っている。一方で、右側に示すような搭乗状態にあるヒューマノイドでは一意的な組み付けがなされるのではなく、機器上における足位置などの搭乗状態に応じた運動の特性の変化が存在し、搭乗した環境の運動に適応したゲイン学習が要求される。

2.3 本研究における探索指標の定義と分類

本章の題目にある「オンライン探索指標」のアプローチを提案するにあたって、はじめに本節では「探索指標」についての定義を行う。さらに、オンライン探索指標を含めた探索指標の種類に関して各々を定義した上で、その分類法および集合を用いた立ち位置の関係、各探索指標の具体的な例を示す。

2.3.1 探索指標の定義

「探索」の基本的な考えは、ある理想的な条件を満たす解を試行錯誤的に状態を変化させながら得ることである。このときの状態変化となる指標（指針）を本研究では「探索指標」と定義する。

探索空間 [78] において現在状態から目標状態に到達するまでのプロセスは数多存在するが、目標状態に至るための指標（指針）を与えることで、理想となる解へと近づく可能性が高まる。探索指標は、現在状態となる変数を探索空間内でどの方向にどの程度動かすか、を表現したものでありベクトルとしても表される。

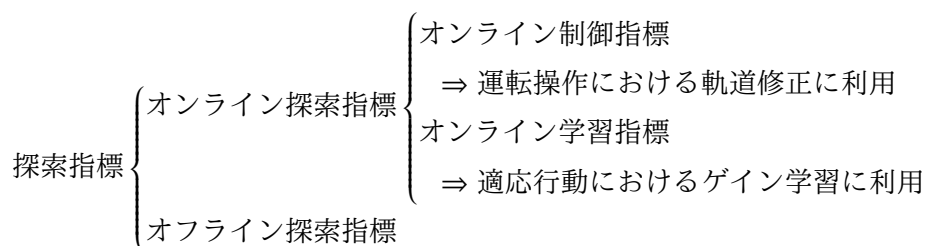
本研究では現在状態を目標状態に近づけることを目的とした探索指標の連続的な結合により、目標状態となる理想的な解を獲得するアプローチを提案する。現在状態や目標状態を示す状態表現は、文献 [78] で定義されるような状態空間上における状態変数の点で表され、ここでの状態変数は主に物理量を対象とする。探索における状態変数を本論文では「探索変数」と定義し、目標状態となる探索変数を「探索目標」として定義する。探索変数は多種多様に設定することが可能であり、本研究では力、角度、情報、トルク、速度などを対象とする。

“探索指標”の定義

試行錯誤的に行われる探索の考えにおいて、現在状態（“探索変数”と定義）を探索空間内でどの方向にどの程度動かせば目標状態（“探索目標”と定義）に近づき得るかを表した指標（指針）となるベクトル。

2.3.2 探索指標の分類

本研究における探索指標の種類とその分類を以下に示す。



探索指標は「オンライン探索指標」と「オフライン探索指標」に二分化され、さらに「オンライン探索指標」は、運転操作における軌道修正に利用する目的の「オンライン制御指標」と、適応行動におけるゲイン学習に利用する目的の「オンライン学習指標」に分けられる。それぞれの探索指標の立ち位置を集合を用いて表した図が図 2.9 である。本研究では図 2.9 で色付けされている「オンライン探索指標」に主に着目し、「オンライン探索指標」が内包する「オンライン制御指標」と「オンライン学習指標」に基づき軌道修正とゲイン学習を実現させる。また、各探索指標の具体的な例を表 2.1 にまとめ、以下でそれぞれの探索指標について説明する。

表 2.1: 分類された各探索指標の具体例

探索指標	オンライン探索指標	オンライン制御指標	<ul style="list-style-type: none"> ・計測接触力と目標力のエラー [48][49][50] ・現在姿勢角と目標姿勢角のエラー [55] ・傾斜角と基準（ゼロ点）のエラー [52] ・ゴールを目標としたナビゲーション情報 [79][80]
		オンライン学習指標	<ul style="list-style-type: none"> ・制御系のゲインに対する評価関数の勾配 [81][82][83] (探索目標におけるポテンシャルが最小となる評価関数を設定)
	オフライン探索指標	<ul style="list-style-type: none"> ・歩行パターン軌道生成 [84] に基づく AutoBalancer [85] ・関節空間における勾配を用いた最適軌道設計 (CHOMP [86][87], STOMP [88]) ・バッチ学習 (オフライン学習) [89][90][91] 	

分類したそれぞれの探索指標における具体的な例を示す。オンライン探索指標については、本研究で着目する搭乗型機器の運転操作と適応行動における課題例に対応している。

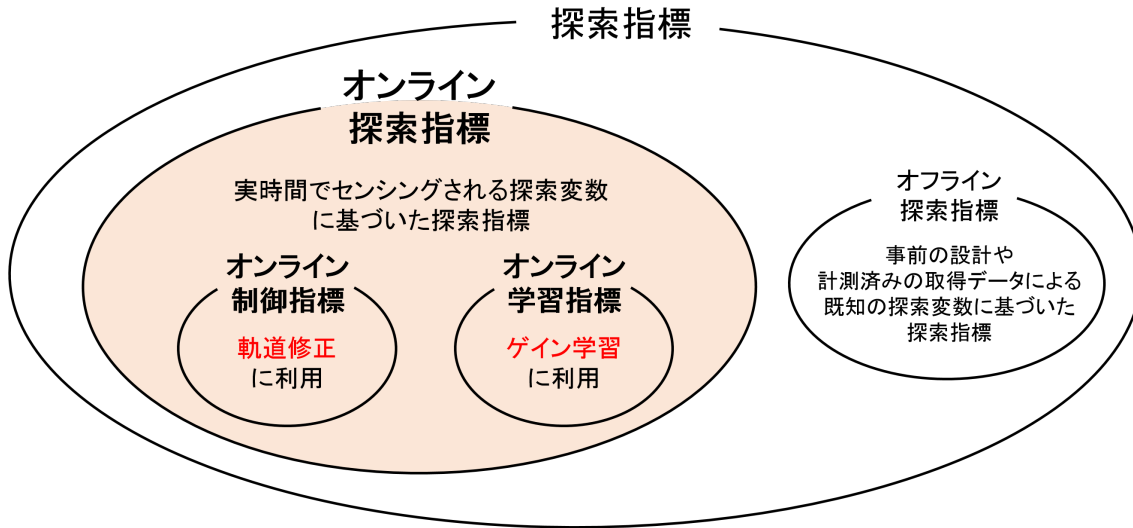


図 2.9: 探索指標の分類

「探索指標」は「オンライン探索指標」と「オフライン探索指標」に分類される。さらに、「オンライン探索指標」は「オンライン制御指標」と「オンライン学習指標」に分類される。色付けされた集合が本研究の柱となる「オンライン探索指標」の立ち位置を表す。「オンライン探索指標」の部分集合である「オンライン制御指標」は軌道修正に、「オンライン学習指標」はゲイン学習に利用される。

オンライン探索指標

探索指標の定義で述べた探索変数の種類が、実時間でセンシングされた状態変数である場合を特に「オンライン探索指標」と呼ぶものとする。したがって、ロボットがオンライン探索指標を利用するためには、センサをハードウェアとして備えていることが必要条件となる。センシングされた状態がもつポテンシャルの空間における基底ベクトルが軌道やゲインを表しており、軌道やゲインの変化に対するポテンシャルの変化率がオンライン探索指標を表す。探索目標の近傍では、軌道やゲインに対して局所的にポテンシャルの変化は発生しない。すなわち、この変化（オンライン探索指標）が発生しないような軌道やゲインを探索することで探索目標を得ることが可能である。搭乗型機器の運転操作や適応行動において探索目標を設定した上で、オンライン探索指標が零ベクトルとなる点の軌道やゲインを探索することが軌道修正やゲイン学習の本質を表している。オンライン探索指標の具体的な表現方法については2.4節、オンライン探索指標を利用した軌道修正とゲイン学習の手法は2.5節にて詳説する。

“オンライン探索指標”の定義

実時間でセンシングされる探索変数に基づいた探索指標。

オンライン制御指標

オンライン制御指標はオンライン探索指標の部分集合（図 2.9）であり，軌道修正において利用される探索指標である。オンライン制御指標は表 2.1 の例に示すように，制御工学の分野で活用されている偏差（エラー）を基にした探索変数と探索目標を結んだ探索指標を表している。制御の基本的な考え方のように，エラーから制御入力を決定して制御対象のプラントに制御入力を与えるプロセスと異なり，目標の軌道を逐次的に修正する軌道修正の考え方に対してオンライン制御指標が適用される。具体的な定式化を 2.4 節，軌道修正への適用方策を 2.5 節で述べる。

オンライン学習指標

オンライン学習指標はオンライン探索指標が内包するもう一つの部分集合（図 2.9）であり，ゲイン学習において利用される探索指標である。オンライン学習指標は，探索目標でのポテンシャルが最小となるような評価関数を設定したときの制御系のゲインに対する探索指標を表している（表 2.1）。探索指標を用いたゲイン学習においては，最急降下法や確率的勾配降下法（Stochastic Gradient Descent: SGD）[92][93][94]などに基づく探索アプローチを講ずる。オンライン制御指標と同様に，定式化を 2.4 節，ゲイン学習への適用方策を 2.5 節で述べる。

オフライン探索指標

実時間でセンシングされる探索変数を利用するオンライン探索指標に対して，「オフライン探索指標」は事前の設計や計測済みの取得データによる既知の探索変数を利用する探索変数と定義する。オフラインの探索変数を用いる例として，二足歩行ロボットの事前の歩行パターン生成法 [84][85] や，動作の軌道計画における最適化手法 [86][87][88]，オフラインでの機械学習手法として知られるバッチ学習 [89][90][91] などが挙げられる（表 2.1）。これらの例においてオフライン探索指標は，事前に計画された軌道のポーズ間の遷移量（関節角

度変位，エンドエフェクタ変位，等）やオフラインでの学習における更新の差分などに相当する。

“オフライン探索指標”の定義

事前の設計や計測済みの取得データによる既知の探索変数に基づいた探索指標。

2.4 オンライン探索指標の定式化

前節で定義した探索指標の概念の内，探索変数が実時間でセンシング・取得される場合をオンライン探索指標と定義した。オンライン探索指標を定式化するにあたって，はじめにこれまで述べてきた探索変数の概念について一般性を考慮した上で表現する。

前節の通り，探索変数は実世界における物理量を表しており，ロボットに搭載されたセンサによりリアルタイムで実環境の状態認知が行われる。搭載するセンサの種類（内界センサや外界センサ）に応じて多種多様な物理量が取得可能である。このときの探索変数を一般化させて表現する場合に共通となる媒介変数（助変数）を考える。オンライン探索指標の定義には「センシング」というワード以外にも、「実時間」というキーワードが含まれている。これはセンサというデバイス固有がもつ特徴の一つであり，実世界で発生した物理現象を実世界における時間とほぼ同時に認知できる基本的な性質を表している。ここで「ほぼ」と表現しているのは，実際の実機環境では電装系の伝達経路における遅延・ロスや，センシングされたデータを処理する計算機上でのレイテンシや離散時間のサンプリングレートの存在があるためである。

以上から，実世界における時間（実時間）に対して，対応する物理現象（物理量として表される状態変数）は一意に決められることから，実時間での時刻を t としたとき，探索変数は t の媒介変数表示で記述できる。また，先に述べた離散時間システムやタイマのカウンタ等により，非負の整数 $n = 0, 1, 2, \dots$ を媒介変数にとり探索変数を表現することもできる。

さて，これまでに述べた探索変数という概念は媒介変数 t や n を用いた状態変数として表すことができる。探索変数には物理量の数値が格納され，次元に限らず多次元の配列・ベクトルを形成することができる。このことから探索変数はベクトルとして表記され， t や n を媒介変数の引数として与えて以下の式で探索変数 $\mathbf{x}(t; n)$ を表現する。

$$\mathbf{x}(t; n) \in \mathbb{R}^N \quad (2.1)$$

ここで、 N は探索変数の次元である。 $\mathbf{x}(t; n)$ の引数内の $;$ は t, n どちらでも媒介変数としてとることが可能なことを表しており、二変数の表記と区別している。探索変数 $\mathbf{x}(t; n)$ には具体的に以下のような実時間センシングに基づいた情報が含まれる。

- 力覚センサにより計測される接触力
- IMU (Inertial Measurement Unit) により計測されるロボットの姿勢角や傾斜角
- SLAM (Simultaneous Localization and Mapping) 等を利用した環境認識点群に基づくナビゲーション情報
- ロボット足裏に搭載された6軸力覚センサにより計測される足部まわりの発生トルク
- 実装した制御系で観測される制御入力
- ロボットの移動スピード (速度や角速度)

探索指標の定義で述べた通り、目標状態に相当する探索変数を「探索目標」と定義し $\mathbf{x}_d \in \mathbb{R}^N$ のベクトルで表す。搭乗型機器の運転操作と適応行動では、探索変数 $\mathbf{x}(t; n)$ を元に軌道の修正やゲインの学習を実行していく。したがって、軌道またはゲインは探索変数と同様の媒介変数 $(t; n)$ 表記が可能であり、同じくベクトルとして軌道またはゲイン $\boldsymbol{\rho}(t; n)$ を以下の式で表す。

$$\boldsymbol{\rho}(t; n) \in \mathbb{R}^M \quad (2.2)$$

M は軌道またはゲインがもつ次元である。

媒介変数 $(t; n)$ におけるオンライン探索指標、オンライン制御指標、オンライン学習指標を表 2.2 の下段に示す式で与える。探索指標の定義より $\mathbf{f}(t; n), \mathbf{f}_c(\mathbf{x}(t; n)), \mathbf{f}_l(\mathbf{x}(t; n))$ は全て軌道またはゲイン $\boldsymbol{\rho}(t; n)$ の次元 M と同次元のベクトルである。

$$\mathbf{f}(t; n) \in \mathbb{R}^M, \mathbf{f}_c(\mathbf{x}(t; n)) \in \mathbb{R}^M, \mathbf{f}_l(\mathbf{x}(t; n)) \in \mathbb{R}^M \quad (2.3)$$

このとき、オンライン探索指標 $\mathbf{f}(t; n)$ は $\mathbf{x}(t; n), \mathbf{x}_d, \boldsymbol{\rho}(t; n)$ を用いて次式で定式化される。

$$\mathbf{f}(t; n) = \begin{cases} \mathbf{f}_c(\mathbf{x}(t; n)) = \mathbf{sgn}[\boldsymbol{\rho}(t; n), \mathbf{x}(t; n)](\mathbf{x}(t; n) - \mathbf{x}_d) \\ \mathbf{f}_l(\mathbf{x}(t; n)) = \left(\frac{\partial J(\mathbf{x}(t; n))}{\partial \boldsymbol{\rho}(t; n)} \right)^\top \end{cases} \quad (2.4)$$

表 2.2: オンライン探索指標の定式化における各パラメータ

$\mathbf{x}(t; n)$	\mathbf{x}_d	$\boldsymbol{\rho}(t; n)$
探索変数	探索目標	軌道またはゲイン
$\mathbf{f}(t; n)$	$\mathbf{f}_c(\mathbf{x}(t; n))$	$\mathbf{f}_l(\mathbf{x}(t; n))$
オンライン探索指標	オンライン制御指標	オンライン学習指標

上記 6 つのパラメータは全てベクトルを表す。 $\mathbf{x}(t; n), \mathbf{x}_d$ は N 次元のベクトル， $\boldsymbol{\rho}(t; n), \mathbf{f}(t; n), \mathbf{f}_c(\mathbf{x}(t; n)), \mathbf{f}_l(\mathbf{x}(t; n))$ は M 次元のベクトルである。各探索指標および軌道やゲインは時刻 t や添字 n を媒介変数にもつ。

式 2.4 の上式がオンライン制御指標の場合であり，下式がオンライン学習指標の場合を表している。オンライン探索指標 $\mathbf{f}(t; n)$ のベクトルの向きは，探索目標 \mathbf{x}_d から離れる方向（斥力方向）に働く。式中の $\text{sgn}[\boldsymbol{\rho}(t; n), \mathbf{x}(t; n)]$ および $J(\mathbf{x}(t; n))$ の関数定義も含めて，以下でそれぞれについて詳説する。

2.4.1 オンライン制御指標の定式化

オンライン制御指標 $\mathbf{f}_c(\mathbf{x}(t; n))$ は $\mathbf{x}(t; n), \mathbf{x}_d, \boldsymbol{\rho}(t; n)$ から以下の式で定義される。

$$\mathbf{f}_c(\mathbf{x}(t; n)) = \text{sgn}[\boldsymbol{\rho}(t; n), \mathbf{x}(t; n)](\mathbf{x}(t; n) - \mathbf{x}_d) \quad (2.5)$$

$\mathbf{f}_c(\mathbf{x}(t; n))$ は，現在の探索変数 $\mathbf{x}(t; n)$ と探索目標 \mathbf{x}_d の偏差 $\mathbf{x}(t; n) - \mathbf{x}_d$ の N 次元ベクトルに対して $M \times N$ 次元の行列 $\text{sgn}[\boldsymbol{\rho}(t; n), \mathbf{x}(t; n)]$ が掛けられることで得られる。行列 $\text{sgn}[\boldsymbol{\rho}(t; n), \mathbf{x}(t; n)]$ は軌道 $\boldsymbol{\rho}(t; n)$ の M 次元ベクトルと $\mathbf{x}(t; n)$ から以下で与えられる。

$$\text{sgn}[\boldsymbol{\rho}(t; n), \mathbf{x}(t; n)] = \begin{pmatrix} \text{sgn}[\rho_1(t; n), x_1(t; n)] & \text{sgn}[\rho_1(t; n), x_2(t; n)] & \dots & \text{sgn}[\rho_1(t; n), x_N(t; n)] \\ \text{sgn}[\rho_2(t; n), x_1(t; n)] & \text{sgn}[\rho_2(t; n), x_2(t; n)] & \dots & \text{sgn}[\rho_2(t; n), x_N(t; n)] \\ \vdots & \vdots & \ddots & \vdots \\ \text{sgn}[\rho_M(t; n), x_1(t; n)] & \text{sgn}[\rho_M(t; n), x_2(t; n)] & \dots & \text{sgn}[\rho_M(t; n), x_N(t; n)] \end{pmatrix} \quad (2.6)$$

ここで，行列を構成する要素 $\text{sgn}[\rho_i(t; n), x_j(t; n)]$ ($i = 1, 2, \dots, M, j = 1, 2, \dots, N$) は $\rho_i(t; n)$ に対する $x_j(t; n)$ の関係に基づき以下の場合分けにより $1, 0, -1$ の各値が振られる。

$$\text{sgn}[\rho_i(t; n), x_j(t; n)] = \begin{cases} 1 & (\rho_i(t; n) \text{ に対して } x_j(t; n) \text{ が広義単調増加の場合}) \\ 0 & (\rho_i(t; n) \text{ に対して } x_j(t; n) \text{ が非依存もしくは不安定の場合}) \\ -1 & (\rho_i(t; n) \text{ に対して } x_j(t; n) \text{ が広義単調減少の場合}) \end{cases} \quad (2.7)$$

各値が振られる条件は $\rho_i(t; n)$ に対する $x_j(t; n)$ の単調性により識別され、 $\rho_i(t; n)$ に対して $x_j(t; n)$ が広義単調増加を示す場合は $\text{sgn}[\rho_i(t; n), x_j(t; n)] = 1$ 、広義単調減少を示す場合は $\text{sgn}[\rho_i(t; n), x_j(t; n)] = -1$ 、そして $\rho_i(t; n)$ の変化に対して $x_j(t; n)$ の相関がない（非依存）もしくは別の $\rho_*(t; n)$ の状態により広義単調増加、広義単調減少、非依存と振動する不安定な場合は $\text{sgn}[\rho_i(t; n), x_j(t; n)] = 0$ を与えるものとして定義する。

具体的に各場合を表す行動例を表 2.3 に示す。例えば、歩行時の足の着地位置の軌道 $\rho_i(t; n)$ が増加するほど、Visual Odometry から推定されるロボットの自己位置 $x_j(t; n)$ は軌道 $\rho_i(t; n)$ の方向へと増加していく。車の操作の例では、ペダルの踏み込み量 $\rho_i(t; n)$ に対して IMU で計測される Yaw 方向の角速度 $x_j(t; n)$ は、車のハンドル角度（別の軌道 $\rho_*(t; n)$ ）の状態に応じて、広義単調増加する場合もあれば広義単調減少する場合もあり、ハンドル角度が 0 に近づけば Yaw まわりの回転を伴うことなく非依存を呈する。障壁等に向かって手先をリーチングさせたときの軌道 $\rho_i(t; n)$ に対しては、リーチング軌道 $\rho_i(t; n)$ が増えるほど壁面から受ける反力の絶対値は大きくなるが、その働く方向は軌道とは逆方向であり、ロボットが力覚センサにより検知する接触力 $x_j(t; n)$ は減少する方向に働く。なお、リーチングの途中過程などで接触が未達成の場合は、軌道 $\rho_i(t; n)$ が増加しても減少しても接触力 $x_j(t; n)$ は変化せず零となることからこの意味でも「広義」として定義している。

以上の単調性の判定は定性的に決定する他、近年取り組まれているロボットの試行的な知覚応答に基づく獲得・判断の手法 [95][96][97][98] を応用することでも得られる。 $\text{sgn}[\rho_i(t; n), x_j(t; n)]$ 関数は式 2.4 で示したオンライン探索指標 $f(t; n)$ の向きを調整するための関数であり、現在と目標の偏差 $\mathbf{x}(t; n) - \mathbf{x}_d$ の符号を切り替える役割を担っている。偏差の符号を変えることで、軌道 $\rho(t; n)$ に対する探索変数 $\mathbf{x}(t; n)$ の反応に応じて適応的に探索目標 \mathbf{x}_d から退く方向にオンライン探索指標 $f(t; n)$ を向けることが可能となる。

2.4.2 オンライン学習指標の定式化

オンライン学習指標 $f_l(\mathbf{x}(t; n))$ は、探索目標においてポテンシャルが最小となる評価関数 $J(\mathbf{x}(t; n))$ のゲイン列 $\rho(t; n)$ に関する微分として以下の M 次元のベクトルとして定義する。

$$f_l(\mathbf{x}(t; n)) = \left(\frac{\partial J(\mathbf{x}(t; n))}{\partial \rho(t; n)} \right)^T \quad (2.8)$$

$J(\mathbf{x}(t; n))$ は探索変数 $\mathbf{x}(t; n)$ を引数にとり、 $\mathbf{x}(t; n)$ が探索目標 \mathbf{x}_d と一致する ($\mathbf{x}(t; n) = \mathbf{x}_d$) とし最小となるスカラー量の評価関数として設定する。したがって、 $J(\mathbf{x}(t; n))$ は次の式を満

表 2.3: $\rho_i(t; n)$ と $x_j(t; n)$ の関係に基づく $\text{sgn}[\rho_i(t; n), x_j(t; n)]$ の割り当てにおける行動例

$\rho_i(t; n)$ に対して $x_j(t; n)$ が 広義単調増加の場合	$\rho_i(t; n)$ に対して $x_j(t; n)$ が 非依存もしくは 不安定の場合	$\rho_i(t; n)$ に対して $x_j(t; n)$ が 広義単調減少の場合
$\text{sgn}[\rho_i(t; n), x_j(t; n)] = 1$	$\text{sgn}[\rho_i(t; n), x_j(t; n)] = 0$	$\text{sgn}[\rho_i(t; n), x_j(t; n)] = -1$
歩行時の足の着地位置の軌道 $\rho_i(t; n)$ に対する, Visual Odometry によりセンシングされる自己位置推定 $x_j(t; n)$	車のペダルの踏み込み量 $\rho_i(t; n)$ に対する, IMU によりセンシングされる Yaw まわりの角速度 $x_j(t; n)$	壁面の法線方向に向かう手先のリーチング軌道 $\rho_i(t; n)$ に対する, 力覚センサによりセンシングされる法線成分の接触力 $x_j(t; n)$

軌道 $\rho_i(t; n)$ の増加方向に対してセンシングされる探索変数 $x_j(t; n)$ が, 増加する方向 (広義単調増加) か, 減少する方向 (広義単調減少) か, 相関のない非依存もしくは状況に応じて変則的に振動する不安定な事象として働くか, により $\text{sgn}[\rho_i(t; n), x_j(t; n)]$ に割り当てられる値が判定される。

たす。

$$\underset{\mathbf{x}(t; n)}{\text{argmin}} J(\mathbf{x}(t; n)) = \mathbf{x}_d \quad (2.9)$$

$J(\mathbf{x}(t; n))$ の設計例としては, 最適化の議論でよく用いられる二次形式による表現などが考えられ, 例えば W を N 次の対称行列として以下のような設定が可能である。

$$J(\mathbf{x}(t; n)) = (\mathbf{x}(t; n) - \mathbf{x}_d)^T W (\mathbf{x}(t; n) - \mathbf{x}_d) \quad (2.10)$$

W は最小化における重み行列 [99] に相当し, 式 2.10 の $J(\mathbf{x}(t; n))$ は $\mathbf{x}(t; n) = \mathbf{x}_d$ で極小かつ最小 (0) となり式 2.9 の条件を満たしている。

さて, このような評価関数 $J(\mathbf{x}(t; n))$ に対してゲインを配列した M 次元ベクトル $\boldsymbol{\rho}(t; n)$ で微分することで, $J(\mathbf{x}(t; n))$ が最小となる目標状態 $\mathbf{x}(t; n) = \mathbf{x}_d$ を満足する理想のゲイン $\boldsymbol{\rho}_*(t; n)$ から遠ざかる方向の M 次元ベクトル $\mathbf{f}_l(\mathbf{x}(t; n))$ を得ることができる。このときの $\mathbf{f}_l(\mathbf{x}(t; n))$ をオンライン学習指標と定義する。オンライン学習指標の向きが目標状態 $\mathbf{x}(t; n) = \mathbf{x}_d$ から退く方向に働くことは, 最急降下法や確率的勾配降下法 [92][93][94] の原理に等しい。

2.4.3 写像とブロック図

本節で新たに定義した探索変数の集合から図 2.9 で示したオンライン探索指標の集合への写像の関係を図 2.10 にまとめる。探索変数 $\mathbf{x}(t; n)$ は式 2.4 によりオンライン探索指標 $\mathbf{f}(t; n)$

へと写される.

$$\mathbf{x}(t;n) \xrightarrow{\text{式 2.4}} \mathbf{f}(t;n) \quad (2.11)$$

特に, オンライン学習指標 $f_l(\mathbf{x}(t;n))$ は設定した評価関数による写像 J によりスカラー場の評価関数空間 $J(\mathbf{x}(t;n))$ を介して以下のように写される.

$$\mathbf{x}(t;n) \xrightarrow{J} J(\mathbf{x}(t;n)) \xrightarrow{\partial/\partial \boldsymbol{\rho}(t;n)} f_l(\mathbf{x}(t;n)) \quad (2.12)$$

集合の関係性より $f_l(\mathbf{x}(t;n)) \in \mathbf{f}(t;n)$ である. 以上のように, オンライン探索指標 $\mathbf{f}(t;n)$ は探索変数 $\mathbf{x}(t;n)$ を元にして生成される概念である.

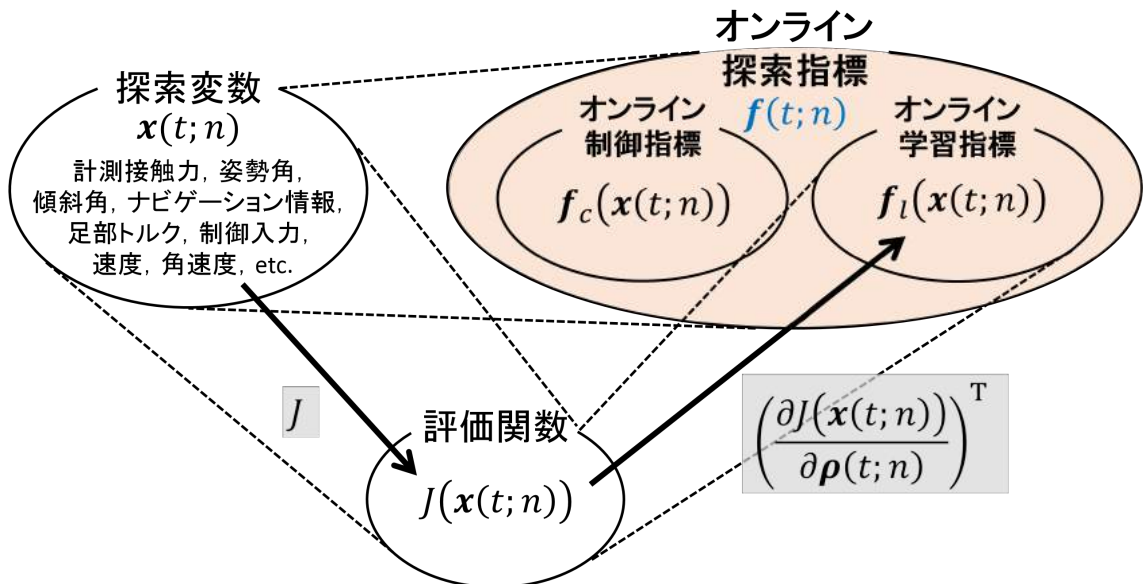


図 2.10: 探索変数 $\mathbf{x}(t;n)$ からオンライン探索指標 $\mathbf{f}(t;n)$ への写像

$\mathbf{x}(t;n)$ から $\mathbf{f}(t;n)$ への写像の関数は式 2.4 で与えられる. オンライン探索指標が内包するオンライン学習指標については, 探索変数から写された評価関数空間を介して生成される.

さらに詳細に計算機上での利用等を想定した場合の構成として, 図 2.11 にオンライン探索指標生成までのプロセスをブロック図で表す. ブロック図 2.11 は式 2.4 に基づくシステム構成となっており, 探索変数 $\mathbf{x}(t;n)$, 探索目標 \mathbf{x}_d , 軌道またはゲイン $\boldsymbol{\rho}(t;n)$ を指定することでオンライン探索指標 $\mathbf{f}(t;n)$ を算出することができる. 算出したオンライン探索指標 $\mathbf{f}(t;n)$ は, 次節の軌道修正ならびにゲイン学習において利用する.

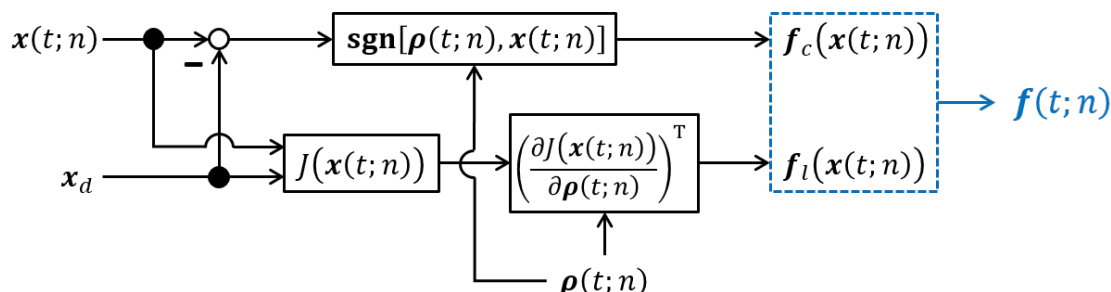


図 2.11: オンライン探索指標の生成におけるブロック図

探索変数 $x(t;n)$ 、探索目標 x_d 、軌道またはゲイン $\rho(t;n)$ を入力として与えることでオンライン探索指標 $f(t;n)$ が生成される。その他、 $\text{sgn}[\rho(t;n), x(t;n)]$ 関数の割り当てや評価関数 $J(x(t;n))$ の設定が必要となる。

2.5 オンライン探索指標を利用した軌道修正とゲイン学習

本節では、2.4 節で定式化したオンライン探索指標 $f(t;n)$ を用いた軌道修正とゲイン学習の方法論について述べる。図 2.9 の探索指標の定義・分類で示したように、搭乗型機器の運転操作における軌道修正に対してはオンライン制御指標 $f_c(x(t;n))$ を利用し、搭乗型機器の適応行動におけるゲイン学習に対してはオンライン学習指標 $f_l(x(t;n))$ を利用する方策であるが、 $f_c(x(t;n)), f_l(x(t;n)) \in f(t;n)$ の関係より共通化されたオンライン探索指標 $f(t;n)$ を用いて軌道修正とゲイン学習の考え方を一般化する。また、一般化された法則に基づいた探索目標に対するアプローチに関しても、軌道修正とゲイン学習におけるそれぞれの動き（修正と学習）を観察し考察する。

2.5.1 オンライン探索指標に基づく軌道とゲインの一般化更新則

オンライン制御指標とオンライン学習指標を一般化したオンライン探索指標 $f(t;n)$ により、軌道またはゲイン $\rho(t;n)$ に対する修正と学習の共通化を図る。「軌道の修正」と「ゲインの学習」はともに、 $\rho(t;n)$ の次ステップでの時刻 $t + \Delta t$ (Δt : サンプル時間) や次番号 $n + 1$ における軌道またはゲイン $\rho(t + \Delta t; n + 1)$ を再帰的に決定していくことであり、ここでの「修正」や「学習」の考え方は「更新」として捉えることができる。したがって、現在の軌道やゲイン $\rho(t;n)$ から次の一手となる軌道やゲイン $\rho(t + \Delta t; n + 1)$ への更新の写像 $\rho(t;n) \rightarrow \rho(t + \Delta t; n + 1)$ が定義できればよく、本研究ではこの写像を“軌道とゲインの一般

化更新則”と呼び、以下の定義式および図 2.12 に示す写像で与えるものとする。

$$\rho(t + \Delta t; n + 1) := \rho(t; n) - Af(t; n) \tag{2.13}$$

$\rho(t; n)$ はオンライン探索指標 $f(t; n)$ と正定値対称行列 $A \in \mathbb{R}^{M \times M}$ を用いて $\rho(t + \Delta t; n + 1)$ へと写される。システム上で構成されるブロック図は図 2.13 のように表される。

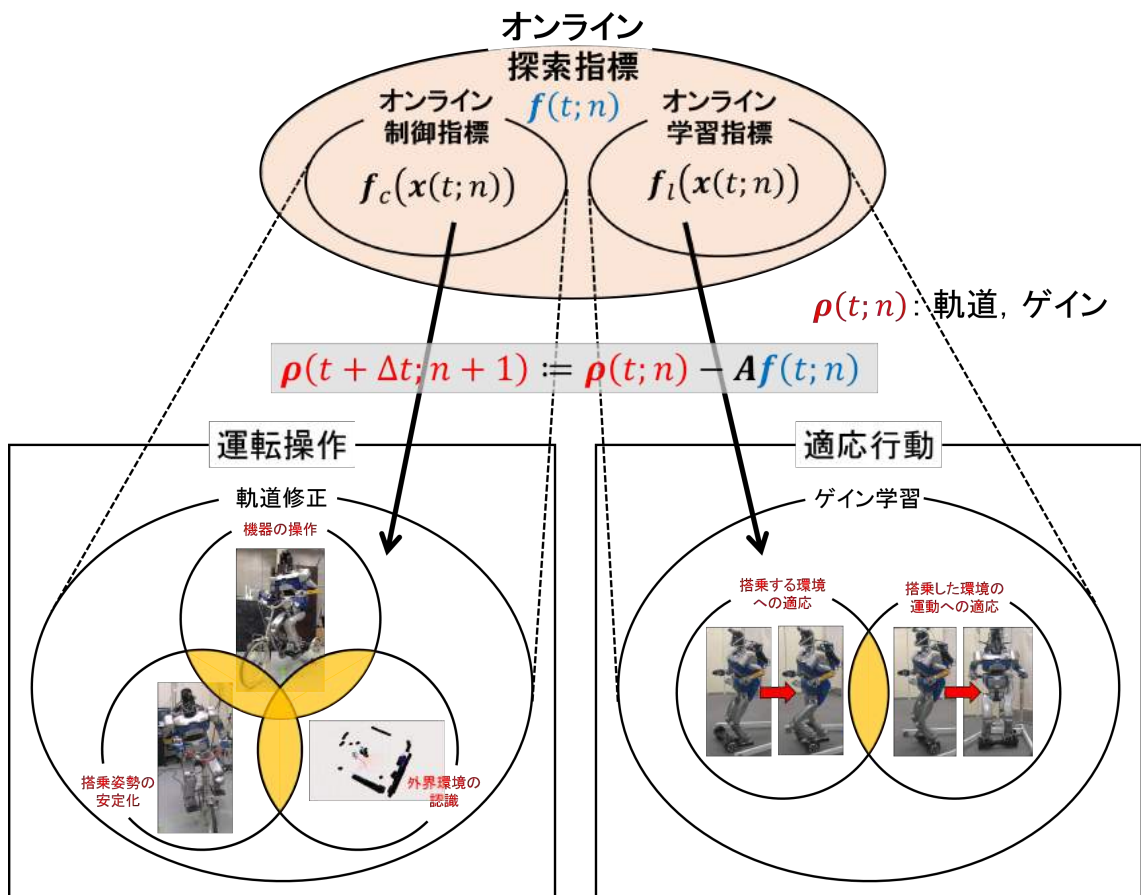


図 2.12: オンライン探索指標に基づく“軌道とゲインの一般化更新則”における写像
 オンライン探索指標 $f(t; n)$ を利用することで軌道またはゲイン $\rho(t; n)$ で共通の一般化された更新の写像 $\rho(t; n) \rightarrow \rho(t + \Delta t; n + 1)$ を定義できる。

式 2.13 はシンプルな形式であるが非常に強い効果を有しており、軌道またはゲイン $\rho(t; n)$ を目標状態 $x(t; n) = x_d$ を満足する最適な軌道またはゲイン $\rho_*(t; n)$ へと逐次的に近づける働きをもつ。この作用は式 2.4 で定義したオンライン探索指標 $f(t; n)$ がもつ探索目標 x_d に対しての後退の性質を利用したものであり、反対に式 2.13 中の $-Af(t; n)$ が探索目標 x_d 対

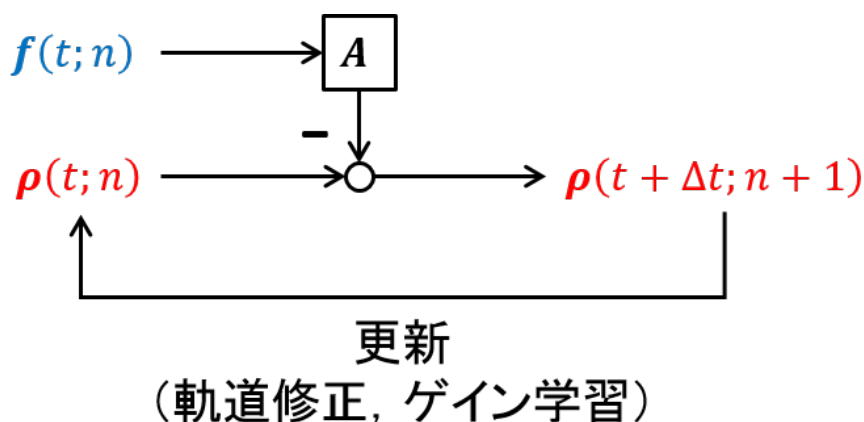


図 2.13: 軌道修正およびゲイン学習で共通する $\rho(t; n)$ 更新システムのブロック図
 $\rho(t; n)$ 更新システムは、式 2.13 で定義した“軌道とゲインの一般化更新則”を構成したブロック図である。正定値対称行列 A を設定した上で、現在状態のオンライン探索指標 $f(t; n)$ と軌道またはゲイン $\rho(t; n)$ を指定することで $\rho(t; n) \rightarrow \rho(t + \Delta t; n + 1)$ の更新が逐次実行される。

しての漸近的な性質を呈している。

制御の観点からみた一般化更新則（式 2.13）の立ち位置は偏差にかかる比例制御として考えることができ、学習の側面からみると最急降下法や確率的勾配降下法 [92][93][94] などの勾配法によるアプローチとして捉えることができる。このときの行列 A は、制御のゲインや勾配法の学習率の役割として振る舞う。以上のように、オンライン探索指標を用いることで制御と学習の両側面から軌道修正やゲイン学習を実行することが可能となる。

2.5.2 軌道修正とゲイン学習による探索目標へのアプローチ

式 2.4 で定義したオンライン探索指標 $f(t; n)$ と“軌道とゲインの一般化更新則”（式 2.13）から、軌道とゲイン $\rho(t; n)$ が修正・学習されることで探索目標 x_d に漸近していく動きを観察する。軌道が修正される動き、およびゲインが学習される動き $\rho(t; n) \rightarrow \rho(t + \Delta t; n + 1)$ はともに、式 2.4 の $f(t; n)$ を式 2.13 に代入することでその更新則が得られる。したがって、軌道修正の更新則は式 2.14、ゲイン学習の更新則は式 2.15 のようにそれぞれ表される。

$$\rho(t + \Delta t; n + 1) := \rho(t; n) - A \operatorname{sgn}[\rho(t; n), \mathbf{x}(t; n)] (\mathbf{x}(t; n) - \mathbf{x}_d) \quad (2.14)$$

$$\rho(t + \Delta t; n + 1) := \rho(t; n) - A \left(\frac{\partial J(\mathbf{x}(t; n))}{\partial \rho(t; n)} \right)^T \quad (2.15)$$

軌道またはゲイン $\rho(t; n)$ を横軸、探索変数 $\mathbf{x}(t; n)$ を縦軸にとる $(\rho(t; n), \mathbf{x}(t; n))$ 空間上におけ

る，探索目標 x_d への軌道修正とゲイン学習の挙動 $\rho(t; n) \rightarrow \rho(t + \Delta t; n + 1)$ ならびにその安定性について以下でそれぞれ述べる．

軌道修正による探索目標への動き

軌道修正の更新則（式 2.14）が移す $\rho(t; n) \rightarrow \rho(t + \Delta t; n + 1)$ の動きを図 2.14 に示す．図 2.14 では $(\rho(t; n), \mathbf{x}(t; n))$ の動きを $(\rho_k, \mathbf{x}_k) = (\rho(k\Delta t; k), \mathbf{x}(k\Delta t; k))$, $k = 0, 1, 2, \dots$ として黒点で描写している．目標状態は (ρ_*, \mathbf{x}_d) として赤点で表しており，この点の ρ_* を探索目標を満足する軌道の解とする．また，この図では ρ に対して \mathbf{x} が広義単調減少する場合 ($\text{sgn}[\rho, \mathbf{x}] = -E$, (E : 単位行列)) を例としている．式 2.14 による軌道修正の軌跡を図 2.14 において矢印で描く．矢印の始点は現在状態 (ρ_k, \mathbf{x}_k) であり，終点は $(\rho_{k+1}, \mathbf{x}_d)$ となる．終点の $(\rho_{k+1}, \mathbf{x}_d)$ と軌道修正後の現在状態 $(\rho_{k+1}, \mathbf{x}_{k+1})$ の差がオンライン制御指標におけるエラーを表現しており，このエラーに応じて修正先の軌道が決定される．このときにどの程度の修正を加えるかの重み付けの係数を「制御指標係数」と呼び，更新式 2.14 より制御指標係数は $(A \text{sgn}[\rho(t; n), \mathbf{x}(t; n)])^+$ で与えられる． $+$ は \cdot の擬似逆行列とする．制御指標係数は図 2.14 の矢印で示した軌跡における ρ に対する \mathbf{x} の傾き度合いに相当し，すべての軌道修正の軌跡で制御指標係数は等しくなる．目標となる軌道の解 ρ_* に近づくほど探索変数 \mathbf{x}_k は探索目標 \mathbf{x}_d に漸近していき，エラー $\mathbf{x}_k - \mathbf{x}_d$ は徐々に収束していく挙動を呈する．以上の原理により，探索変数と探索目標のエラーを小さくすることで逐次的に目標の軌道 ρ_* へと漸近させるアプローチが軌道修正である．

軌道修正における安定性

軌道修正の最終目標は，探索変数 $\mathbf{x}(t; n)$ が探索目標 \mathbf{x}_d に近づくような理想的な軌道 $\rho(t; n) = \rho_*$ を得ることである．本項では，軌道/ゲイン $\rho(t; n)$ の変化量が漸化的に単調減少するときの安定性を“軌道/ゲインの安定性”と定義し，以下に示す命題を安定性が保たれるための条件とする．

命題

探索変数 $\mathbf{x}(t; n)$ が探索目標 \mathbf{x}_d に収束

⇒ 軌道/ゲイン $\rho(t; n)$ の変化量が漸化的に単調減少（軌道/ゲインの安定性）

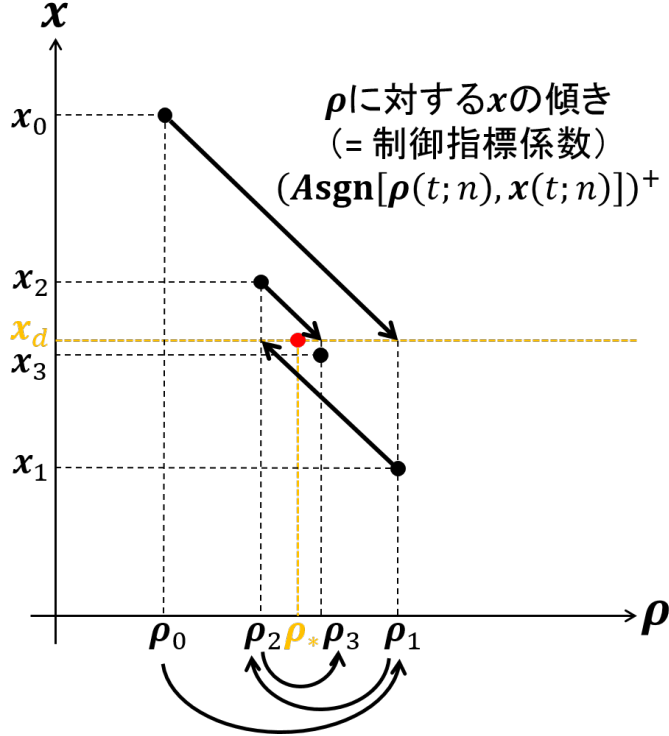


図 2.14: 探索目標に向かう軌道修正の動き

目標状態 (ρ_*, x_d) を赤点，現在状態 (ρ_k, x_k) を黒点，軌道修正による更新の軌跡（式 2.14）を矢印で表現する．探索変数 x_k と探索目標 x_d のエラーに応じて修正先の軌道が決定される．探索目標 x_d まわりではエラーが零に近づき，目標となる軌道 ρ_* に収束する．更新の軌跡における ρ に対しての x の傾きを制御指標係数 $(\text{Asgn}[\rho, x])^+$ として，軌道修正において調整されるパラメータとなる．

命題の仮定「探索変数 $x(t; n)$ が探索目標 x_d に収束」を満足する条件式は漸化的に以下のように表される．

$$\|x_{k+1} - x_d\| < \|x_k - x_d\| \quad (2.16)$$

ここで，前述の議論と同様に $(\rho_k, x_k) = (\rho(k\Delta t; k), x(k\Delta t; k))$ ， $k = 0, 1, 2, \dots$ とする．

はじめに，式 2.16 が成り立つときに，図 2.14 での軌道 ρ の変化量ノルム $\|\rho_{k+1} - \rho_k\|$ に関して，次ステップにおける変化量ノルム $\|\rho_{k+2} - \rho_{k+1}\|$ との大小を比較する．変化量ノルム $\|\rho_{k+1} - \rho_k\|$ は更新式 2.14 に基づいて次式で整理される．

$$\begin{aligned} \|\rho_{k+1} - \rho_k\| &= \|\rho_k - \text{Asgn}[\rho_k, x_k](x_k - x_d) - \rho_k\| \\ &= \|\text{Asgn}[\rho_k, x_k](x_k - x_d)\| \end{aligned} \quad (2.17)$$

ここで，行列 $\text{Asgn}[\rho_k, x_k] \in \mathbb{R}^{M \times N}$ のノルムとベクトル $(x_k - x_d) \in \mathbb{R}^N$ のノルムの性質から

以下の不等式が成立する.

$$\begin{aligned} & \| -\text{Asgn}[\boldsymbol{\rho}_k, \mathbf{x}_k](\mathbf{x}_k - \mathbf{x}_d) \| \leq \| -\text{Asgn}[\boldsymbol{\rho}_k, \mathbf{x}_k] \| \| \mathbf{x}_k - \mathbf{x}_d \| \\ \Leftrightarrow & \| \boldsymbol{\rho}_{k+1} - \boldsymbol{\rho}_k \| \leq \| -\text{Asgn}[\boldsymbol{\rho}_k, \mathbf{x}_k] \| \| \mathbf{x}_k - \mathbf{x}_d \| \end{aligned} \quad (2.18)$$

同様に, 次ステップでの変化量ノルム $\| \boldsymbol{\rho}_{k+2} - \boldsymbol{\rho}_{k+1} \|$ は以下の関係となる.

$$\| \boldsymbol{\rho}_{k+2} - \boldsymbol{\rho}_{k+1} \| \leq \| -\text{Asgn}[\boldsymbol{\rho}_{k+1}, \mathbf{x}_{k+1}] \| \| \mathbf{x}_{k+1} - \mathbf{x}_d \| \quad (2.19)$$

ここで, 現ステップ (現時刻) k と次ステップ (次時刻) $k+1$ の間では, 軌道/ゲイン $\boldsymbol{\rho}$ に対する \mathbf{x} の単調性の判別 (式 2.7) は既定であり変化しないと仮定し, ステップ (時刻) の変化に対する単調性の割り当ての不変性から次式が成り立つものとする.

$$\text{sgn}[\boldsymbol{\rho}_k, \mathbf{x}_k] = \text{sgn}[\boldsymbol{\rho}_{k+1}, \mathbf{x}_{k+1}] \quad (2.20)$$

式 2.16 と式 2.20 から以下の関係式が得られる.

$$\| -\text{Asgn}[\boldsymbol{\rho}_{k+1}, \mathbf{x}_{k+1}] \| \| \mathbf{x}_{k+1} - \mathbf{x}_d \| < \| -\text{Asgn}[\boldsymbol{\rho}_k, \mathbf{x}_k] \| \| \mathbf{x}_k - \mathbf{x}_d \| \quad (2.21)$$

得られた式 2.18, 式 2.19, 式 2.21 の関係から $\| \boldsymbol{\rho}_{k+1} - \boldsymbol{\rho}_k \|$ と $\| \boldsymbol{\rho}_{k+2} - \boldsymbol{\rho}_{k+1} \|$ の大きさを判定すると, 必ずしも $\| \boldsymbol{\rho}_{k+2} - \boldsymbol{\rho}_{k+1} \| < \| \boldsymbol{\rho}_{k+1} - \boldsymbol{\rho}_k \|$ なる単調減少の条件を満足するとは限らない結論が得られる. これは式 2.18 のように探索変数と探索目標のエラー $\mathbf{x}_k - \mathbf{x}_d$ をベクトル量として扱っていることで不等式制約を受けることに起因している.

上記の議論に基づき, 式 2.16 の条件式において, 以下のような単一の探索変数と探索目標のエラー値 (スカラー) $\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j}$ に関する強めた制約条件を与える. ここで, $\{\mathbf{v}\}_{\text{row}=j}$ はベクトル \mathbf{v} の j 行目の値を返す関数とする.

$$|\{\mathbf{x}_{k+1} - \mathbf{x}_d\}_{\text{row}=j}| < |\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j}|, \quad \forall j = 1, 2, \dots, N \quad (2.22)$$

軌道 $\boldsymbol{\rho}$ の変化量 $\boldsymbol{\rho}_{k+1} - \boldsymbol{\rho}_k$ の i 行目の成分 $\{\boldsymbol{\rho}_{k+1} - \boldsymbol{\rho}_k\}_{\text{row}=i}$ ($i = 1, 2, \dots, M$) について書き下す

と次のように記述できる.

$$\begin{aligned}
\{\rho_{k+1} - \rho_k\}_{\text{row}=i} &= \{-\text{Asgn}[\rho_k, \mathbf{x}_k](\mathbf{x}_k - \mathbf{x}_d)\}_{\text{row}=i} \\
&= \{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i}(\mathbf{x}_k - \mathbf{x}_d) \\
&= \{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i, \text{col}=1}\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=1} \\
&\quad + \{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i, \text{col}=2}\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=2} \\
&\quad + \cdots + \{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i, \text{col}=j}\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j} \\
&\quad + \cdots + \{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i, \text{col}=N}\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=N} \\
&= \sum_{j=1}^N \{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i, \text{col}=j}\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j} \tag{2.23}
\end{aligned}$$

$\{V\}_{\text{row}=i, \text{col}=j}$ は行列 V の i 行 j 列目の要素を返す関数である.

ここで, 指定された軌道 $\{\rho_{k+1} - \rho_k\}_{\text{row}=i}$ に対して, ある探索変数と探索目標のエラー $\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j}$ のみが依存 (他の探索変数には非依存) して変化する場合, すなわち

$$\{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i, \text{col} \neq j} = 0 \tag{2.24}$$

の条件を満たす場合には, $\{\rho_{k+1} - \rho_k\}_{\text{row}=i}$ は次式で表される.

$$\{\rho_{k+1} - \rho_k\}_{\text{row}=i} = \{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i, \text{col}=j}\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j} \tag{2.25}$$

式 2.25 より軌道 $\{\rho\}_{\text{row}=i}$ の変化量 $|\{\rho_{k+1} - \rho_k\}_{\text{row}=i}|$ は, $\{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i, \text{col}=j}$ および $\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j}$ がスカラーであることに留意すれば, 以下のように書ける.

$$\begin{aligned}
|\{\rho_{k+1} - \rho_k\}_{\text{row}=i}| &= | \{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i, \text{col}=j}\{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j} | \\
&= | \{-\text{Asgn}[\rho_k, \mathbf{x}_k]\}_{\text{row}=i, \text{col}=j} | | \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j} | \tag{2.26}
\end{aligned}$$

同様に, 次ステップでの変化量 $|\{\rho_{k+2} - \rho_{k+1}\}_{\text{row}=i}|$ は以下となる.

$$|\{\rho_{k+2} - \rho_{k+1}\}_{\text{row}=i}| = | \{-\text{Asgn}[\rho_{k+1}, \mathbf{x}_{k+1}]\}_{\text{row}=i, \text{col}=j} | | \{\mathbf{x}_{k+1} - \mathbf{x}_d\}_{\text{row}=j} | \tag{2.27}$$

制約条件式 2.22, 式 2.26, 式 2.27, 式 2.20 より $|\{\rho_{k+1} - \rho_k\}_{\text{row}=i}|$ と $|\{\rho_{k+2} - \rho_{k+1}\}_{\text{row}=i}|$ の大小関係は以下で結論付けられる.

$$|\{\rho_{k+2} - \rho_{k+1}\}_{\text{row}=i}| < |\{\rho_{k+1} - \rho_k\}_{\text{row}=i}| \tag{2.28}$$

式 2.28 は、軌道 $\{\rho\}_{\text{row}=i}$ の変化量 $|\{\rho_{k+1} - \rho_k\}_{\text{row}=i}|$ が漸化的に単調減少しながら 0 に向かうことを意味しており、最終的に変化量が 0 近傍で抑えられるような“軌道/ゲインの安定性”を満足する。

制約条件式 2.22，および軌道 $\{\rho\}_{\text{row}=i}$ と探索変数 $\{x\}_{\text{row}=j}$ の依存関係に関する条件式 2.24 を満たす場合，命題の結論「軌道/ゲイン $\rho(t; n)$ の変化量が漸化的に単調減少（軌道/ゲインの安定性）」が導かれる。

ゲイン学習による探索目標への動き

ゲイン学習を用いた探索目標への漸近的なアプローチでは，図 2.10 で示した評価関数空間への媒介写像 J により写される $(\rho(t; n), J(x(t; n)))$ 空間において挙動を観察する。 $(\rho(t; n), x(t; n))$ 空間および写像変換後の $(\rho(t; n), J(x(t; n)))$ 空間でのゲイン学習の動きを図 2.15 に示す。図 2.15 の $(\rho(t; n), J(x(t; n)))$ 空間で描写される $J(x_k)$ の集合を青い曲線で表す。 $J(x_k)$ は式 2.9 の評価関数の定義から $x_k = x_d$ で最小となり，最小値 $J(x_d)$ をとるときのゲイン ρ_* が目標となる解である。解となる点 $(\rho_*, J(x_d))$ を $(\rho, J(x))$ 空間において赤点で示し，現在状態 $(\rho_k, J(x_k))$ を黒点で表す。 $(\rho, J(x))$ 空間で描かれる矢印がゲイン学習の更新則（式 2.15）による軌跡であり，この軌跡は始点を $(\rho_k, J(x_k))$ とする傾き $\left(\frac{\partial J(x_k)}{\partial \rho_k}\right)^T$ のベクトルとして表され，その向きは $-A\left(\frac{\partial J(x_k)}{\partial \rho_k}\right)^T$ で決められる。ベクトルの向きは評価関数 $J(x_k)$ が減少する方向に働き，傾きの程度 $\left(\frac{\partial J(x_k)}{\partial \rho_k}\right)^T$ に応じてゲイン ρ_k が更新される。評価関数の減少する方向ならびに傾きの度合いにより，逐次的に $J(x_k)$ の最小化すなわち $J(x_d)$ へと漸近される。 $J(x_d)$ における傾きは零であることから，目標となるゲイン ρ_* まわりでは安定性が保たれている。

ゲイン学習における安定性

ゲイン学習に関する安定性の議論として，本項では式 2.10 で示した二次形式の評価関数 J における安定性を考える。ここで，評価関数 J の重み行列 $W \in \mathbb{R}^{N \times N}$ は正定値対称行列を表し，各探索変数とその探索目標の間のエラー $x_k - x_d$ の大きさ（二乗値）に重み付けされ

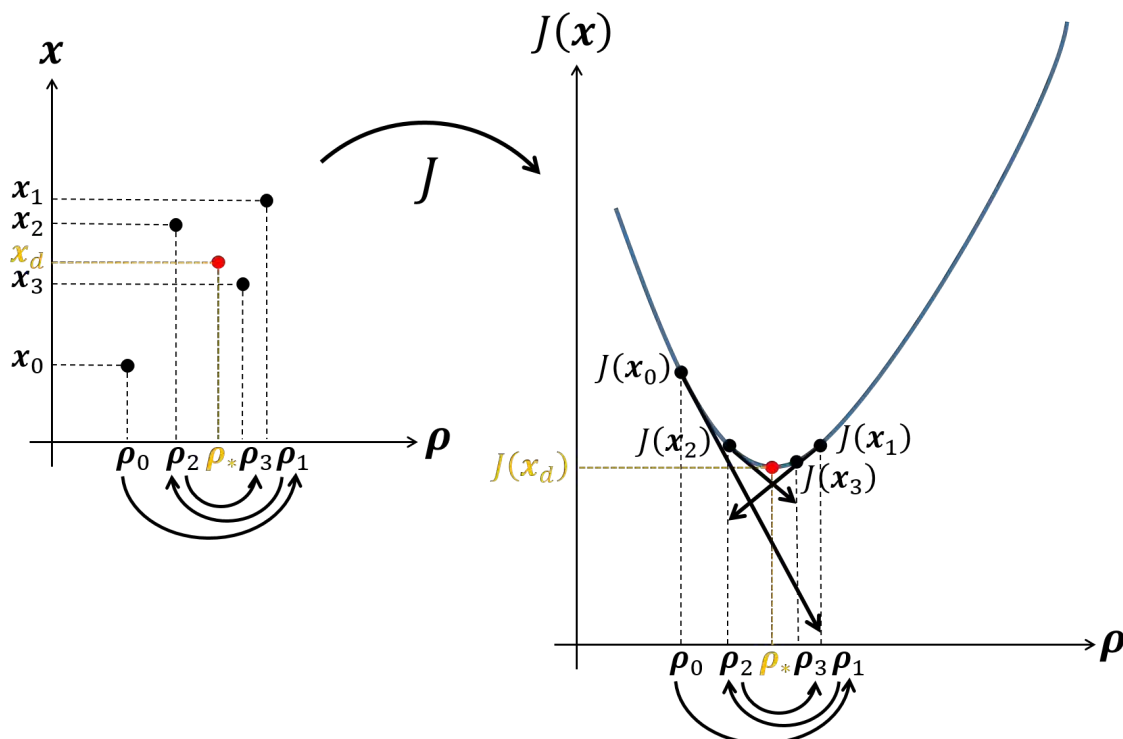


図 2.15: 探索目標に向かうゲイン学習の動き

設定した評価関数の写像 J により写される $(\rho, J(x))$ 空間でゲイン学習の動きを観察する. $(\rho, J(x))$ 空間における目標状態 $(\rho_*, J(x_d))$ を赤点, 現在状態 $(\rho_k, J(x_k))$ を黒点, ゲイン学習による更新の軌跡 (式 2.15) を矢印で表現する. ρ_k に対する $J(x_k)$ の傾き $\left(\frac{\partial J(x_k)}{\partial \rho_k}\right)^T$ により修正先のゲインが決定される. 式 2.9 の評価関数定義から $J(x_d)$ は最小値をとり, 目標となるゲインの解 ρ_* まわりでは傾きが零となり収束に向かう.

た総和となるよう, 重み行列 W は以下のような正定値をとる対角行列として設定する.

$$W = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_N \end{pmatrix} = \text{diag}(w_1, w_2, \dots, w_N), \quad w_1, w_2, \dots, w_N > 0 \quad (2.29)$$

このとき評価関数 $J(\mathbf{x}_k)$ は以下のように整理される.

$$\begin{aligned}
J(\mathbf{x}_k) &= (\mathbf{x}_k - \mathbf{x}_d)^\top W (\mathbf{x}_k - \mathbf{x}_d) \\
&= \begin{pmatrix} \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=1} \\ \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=2} \\ \vdots \\ \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=N} \end{pmatrix}^\top \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_N \end{pmatrix} \begin{pmatrix} \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=1} \\ \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=2} \\ \vdots \\ \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=N} \end{pmatrix} \\
&= w_1 \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=1}^2 + w_2 \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=2}^2 + \dots + w_N \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=N}^2 \\
&= \sum_{j=1}^N w_j \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j}^2 \tag{2.30}
\end{aligned}$$

軌道修正における安定性の議論で用いた制約条件式 2.22 の両辺を二乗した後、両辺に $w_j > 0$ を掛けることで次の不等式が成り立つ.

$$w_j \{\mathbf{x}_{k+1} - \mathbf{x}_d\}_{\text{row}=j}^2 < w_j \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j}^2 \tag{2.31}$$

$\forall j = 1, 2, \dots, N$ において式 2.31 は常に成立することから、総和をとることで以下の不等式についても成り立つ.

$$\sum_{j=1}^N w_j \{\mathbf{x}_{k+1} - \mathbf{x}_d\}_{\text{row}=j}^2 < \sum_{j=1}^N w_j \{\mathbf{x}_k - \mathbf{x}_d\}_{\text{row}=j}^2 \tag{2.32}$$

式 2.30, 式 2.32 より, $J(\mathbf{x}_k)$ と次ステップにおける $J(\mathbf{x}_{k+1})$ の大小関係は以下で結論付けられる.

$$J(\mathbf{x}_{k+1}) < J(\mathbf{x}_k) \tag{2.33}$$

式 2.33 は探索変数 \mathbf{x}_k が更新される度に漸化的に評価関数 $J(\mathbf{x}_k)$ が最小化されていくことを表している. 評価関数 $J(\mathbf{x}_k)$ は図 2.15 でも示されている目標状態 $(\boldsymbol{\rho}_*, J(\mathbf{x}_d))$ の評価関数 $J(\mathbf{x}_d)$ で下に有界として抑えられている ($J(\mathbf{x}_k) \geq J(\mathbf{x}_d)$) ことから, 式 2.33 で示した \mathbf{x}_k に対する $J(\mathbf{x}_k)$ の単調減少の性質と併せて, 最終的に $\mathbf{x}_k = \mathbf{x}_d$ なるゲインの収束値 $\boldsymbol{\rho}_*$ への安定性が保たれる.

以上の議論から, 命題の仮定条件 (制約条件式 2.22) および評価関数の重み行列 W を式 2.29 のようにとるとき, ゲイン学習における安定性の命題の結論「軌道/ゲイン $\boldsymbol{\rho}(t; n)$ の変化量が漸化的に単調減少 (軌道/ゲインの安定性)」が示される.

2.6 搭乗型機器の運転操作と適応行動に対するオンライン探索指標の アプローチ

本研究の狙いとなるヒューマノイドにおける搭乗型機器の運転操作と適応行動に向けて、これまでに述べたオンライン探索指標の考え方を出発点とする全体のアプローチについて本節で論じる。オンライン探索指標の概念を源流として、搭乗型機器の運転操作と適応行動を最終的な着地点としたときの全体の流れは次のようになる。

オンライン探索指標 $\xrightarrow{2.5 \text{ 節}}$ 軌道修正とゲイン学習 $\xrightarrow{2.2 \text{ 節}}$ 運転操作と適応行動

オンライン探索指標の考えから運転操作と適応行動の実現までにおいては、軌道修正とゲイン学習のアプローチを中間に媒介する。オンライン探索指標から軌道修正とゲイン学習までのアプローチについては、2.5 節で述べた一般化された法則の適用により流れが生み出される。軌道修正とゲイン学習から運転操作と適応行動までのアプローチについては、2.2 節で述べた必要性の議論に基づく流れが生み出される。本節では、これら2つの流れを連結することで生まれるオンライン探索指標に基づいた搭乗型機器の運転操作と適応行動への統合的なアプローチに関して詳説する。

全体のシステム構成を図 2.16 に示す。図 2.16 は、図 2.11 のオンライン探索指標生成までのブロック図と、図 2.13 のオンライン探索指標を利用した軌道修正とゲイン学習の更新システムを統合したブロック図である。ブロック図 2.16 は左端の探索変数 $\mathbf{x}(t; n)$ および探索目標 \mathbf{x}_d を入力として軌道またはゲイン $\boldsymbol{\rho}(t; n)$ をリアルタイムに更新していく。探索変数 $\mathbf{x}(t; n)$ と探索目標 \mathbf{x}_d はオンライン探索指標 $\mathbf{f}(t; n)$ の元となる最小構成単位であり、この要素を指定することが基本となる。 $\mathbf{x}(t; n), \mathbf{x}_d$ は言うなれば、ヒューマノイドにおける行動実現の際に我々人間がこの部分（物理量）をこのようにしたいと、人間による観察・判断に基づき決定されうる所望のパラメータに相当する。この基本要素に加えて、ヒューマノイドが何をどうするべきかについて行動で表現される要素が軌道やゲイン $\boldsymbol{\rho}(t; n)$ となる。

したがって、全体システム構成図 2.16 を稼動させるためには最小限として、 $\mathbf{x}(t; n)$ （および \mathbf{x}_d ）と $\boldsymbol{\rho}(t; n)$ を与える必要がある。また、付加的な要素として $\text{sgn}[\boldsymbol{\rho}(t; n), \mathbf{x}(t; n)]$ 関数や評価関数 $J(\mathbf{x}(t; n))$ 、軌道とゲインの一般化更新則の重み付け A の設定が要求される。

以上のことから、搭乗型機器の運転操作と適応行動においては、オンライン探索指標 $\mathbf{f}(t; n)$ の元となる探索変数 $\mathbf{x}(t; n)$ と、更新される軌道/ゲイン $\boldsymbol{\rho}(t; n)$ を決めればよく、運転操作（機器の操作・搭乗姿勢の安定化・外界環境の認識）と適応行動（搭乗する環境への適応・搭乗

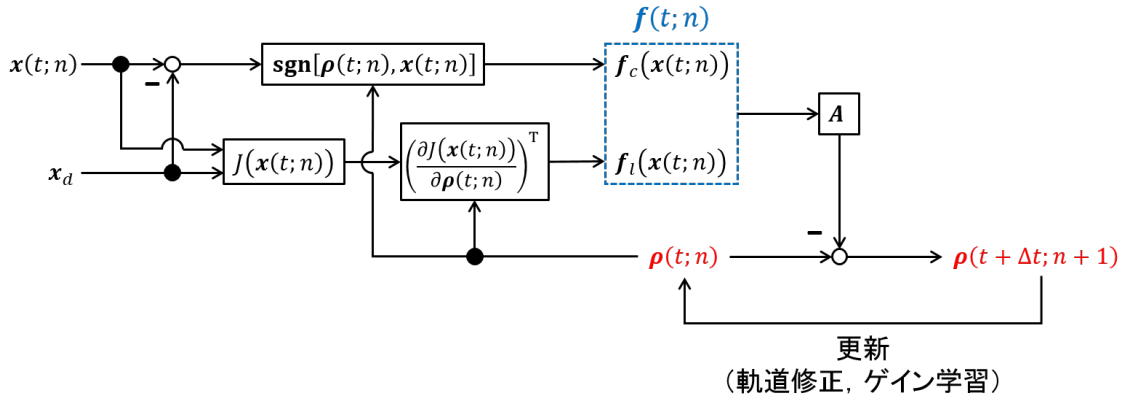


図 2.16: オンライン探索指標に基づく全体のシステム構成図

システム構成図は，オンライン探索指標生成のブロック図 2.11 と軌道/ゲインに対する更新システムのブロック図 2.13 の結合により組成される．オンライン探索指標に基づく搭乗型機器の運転操作と適応行動においては，システム上の入力 $x(t;n)$ （ならびに x_d ）と更新される軌道/ゲイン $\rho(t;n)$ を指定することが基本条件となる．

した環境の運動への適応)のそれぞれについて，表 2.4 に示すような所望の情報を本研究で指定する．表 2.4 は左列に目的とする運転操作と適応行動の 5 要素を示し，上行に必須要素である $x(t;n)$ ならびに $\rho(t;n)$ をとり，各戦略で与える適用内容をそれぞれ記している．以降の各章では表 2.4 に示す構成要素に基づいた上で，搭乗型機器の運転操作と適応行動に対してのアプローチを論じる．

2.7 本章のまとめ

本章では，本論文が目的とする「ヒューマノイドにおける搭乗型機器の運転操作と適応行動」を実現するためのアプローチについて述べた．搭乗型機器の「運転操作と適応行動」における「軌道修正とゲイン学習」の 2 つの方策の重要性を論じた上で，本研究がとるアプローチとして「オンライン探索指標」の考えを提案した．オンライン探索指標を定義するにあたって，はじめに本研究における「探索指標」の意義と派生する探索指標の種類・分類を定めた．次に定義に基づいたオンライン探索指標の定式化を示し，計算処理が行える形式でオンライン探索指標の生成におけるシステムを構成した．さらに定式化したオンライン探索指標を利用して，「運転操作における軌道修正」と「適応行動におけるゲイン学習」に対して共通に適用可能な“軌道とゲインの一般化更新則”を提案した上で，「軌道修正」と「ゲイン学習」による目標状態への各々の振る舞いを述べた．最後に「オンライン探索指標」，「軌道

表 2.4: 搭乗型機器の運転操作と適応行動に対して本研究で与える $x(t; n)$ と $\rho(t; n)$

		オンライン探索指標 $f(t; n)$ の元となる探索変数 $x(t; n)$	更新される 軌道/ゲイン $\rho(t; n)$
運転操作	機器の操作	ヒューマノイドと搭乗型機器の間の接触力（内力）	エンドエフェクタ（操作における効果器）の目標位置
	搭乗姿勢の安定化	ヒューマノイドの姿勢角や傾斜角	体幹の目標姿勢や重心の位置
	外界環境の認識	ゴールまでの二次元のナビゲーション情報	搭乗型機器への目標操作量（ハンドル角度やクランク角度）
適応行動	搭乗する環境への適応	搭乗する環境で発生する足部のトルクとプラントに与える制御入力	足部のトルク制御器のゲイン
	搭乗した環境の運動への適応	搭乗した環境における運動（速度や角速度）	運動に対する制御器のゲイン

本研究で対象とする5つの戦略（機器の操作・搭乗姿勢の安定化・外界環境の認識・搭乗する環境への適応・搭乗した環境の運動への適応）を実現するための最小構成の組 $(x(t; n), \rho(t; n))$ を示す。以降の各章においてそれぞれの適用要素について詳説する。

修正とゲイン学習」, 「運転操作と適応行動」の3つの流れを整理し, 源流となるオンライン探索指標の概念から, 最終的に達成したい搭乗型機器の運転操作と適応行動に至るまでのアプローチをまとめた。

また, 本章では以上の内容を集合と写像による概念を用いて表現した。本章で論じた「探索指標」や「オンライン探索指標」の考え, 「運転操作における軌道修正」や「適応行動におけるゲイン学習」, また「軌道修正」や「ゲイン学習」が解く「機器の操作・搭乗姿勢の安定化・外界環境の認識・搭乗する環境への適応・搭乗した環境の運動への適応」をそれぞれ集合で表したときの各関係および写像をまとめた全体像を図 2.17 に示す。本論文の軸となるオンライン探索指標 $f(t; n)$ の集合は探索変数 $x(t; n)$ を元に写され, $f(t; n)$ の集合を軌道修正やゲイン学習で一般化した更新 $\rho(t; n) \rightarrow \rho(t + \Delta t; n + 1)$ の集合へと写す写像が, 式 2.13 で定義した“軌道とゲインの一般化更新則”に相当している。

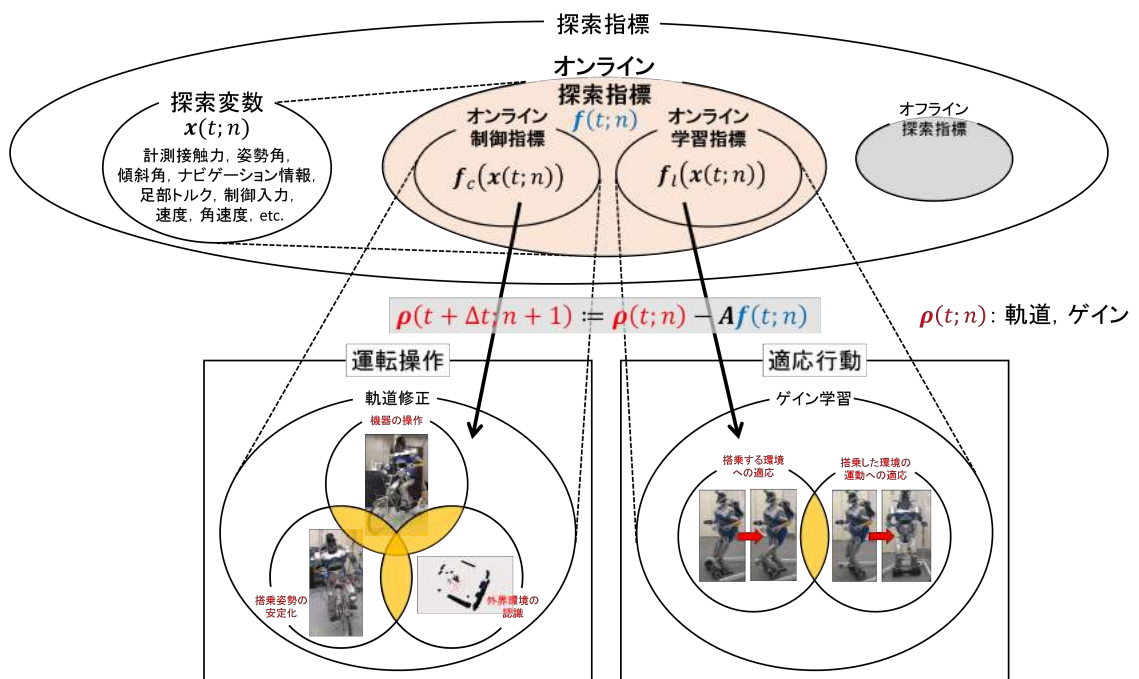


図 2.17: 本章で述べたアプローチにおける集合の関係と写像の表現

本章で焦点を当てたオンライン探索指標 $f(t;n)$ (赤色で塗られた集合) は、探索指標という枠組みの集合における部分集合であり探索変数 $x(t;n)$ から形成される。“軌道とゲインの一般化更新則”(式 2.13) による写像によって、オンライン探索指標に基づく運転操作の軌道修正と適応行動のゲイン学習へと射影がなされる ($\rho(t;n) \rightarrow \rho(t + \Delta t;n + 1)$)。一般化された写像を用いることで、軌道修正による“機器の操作・搭乗姿勢の安定化・外界環境の認識”と、ゲイン学習による“搭乗する環境への適応・搭乗した環境の運動への適応”を解く。

第3章

搭乗型機器の操作に対する軌道修正

3.1 はじめに

本章では、オンライン探索指標に基づく軌道修正を適用することで搭乗型機器の操作の実現を目標とする。ヒューマノイドによる搭乗型機器の操作では、ロボットの腕のみを用いたマニピュレーション戦略・物体操作戦略と異なり、ロボット脚部も協調させた複合的な全身操作行動が要求される。全身操作行動においては、ヒューマノイドと操作の対象となる搭乗型機器との複数点での環境接触が課題であり、ヒューマノイドと搭乗型機器の間に働く拘束性について考慮する必要がある。この拘束性を視野に入れた上で、第2章で定義した探索指標の考えを軸にしたヒューマノイドのマニピュレータにおける軌道計画法および軌道修正法に関して議論していく。探索指標に基づき提案した軌道の制御法の有用性を示すべく、ヒューマノイド実機による双腕・双脚を用いた三輪車の運転操作行動を最終的に実証する。

3.2節では、搭乗型機器の操作の例として三輪車とヒューマノイドの間に介在する閉ループ系の接触拘束について、その定義と分類を踏まえた上で解くべき課題を明確化する。3.3節では、ヒューマノイドが脚腕（双腕と双脚）を協調させて三輪車の操作行動を実現するために必要な戦略を計画し、第2章で論じた探索指標の概念を用いた行動戦略の方法論を述べる。3.4節では、双腕と双脚のマニピュレータに対して“能動機能”と“受動機能”の2つの制御機能を与え、双腕と双脚で共通の探索指標（オフライン探索指標とオンライン探索指標）を用いた軌道制御法を提案する。3.5節では、提案した制御法に基づくヒューマノイドによる三輪車操作について実践を意識した実機検証を示す。

3.2 ヒューマノイドにおける閉ループ系の接触拘束

本節では、ヒューマノイドが環境（操作対象物体も含む）と接触する場合に働く閉ループ系の拘束の構造について着眼する。閉ループ系を呈する接触拘束の種類・分類の定義に基づき、三輪車を例とした搭乗型機器の操作においてヒューマノイドに適用される閉ループ系の拘束について考える。

3.2.1 ヒューマノイド身体と接触対象の間に働く閉ループ系拘束

ヒューマノイドは我々人間が生活する空間において共存し、実世界に働きかける存在となることが期待される。特に身の回りにある道具や物体を操作する研究は過去から現在に至るまで幅広く取り組まれている [100][101][96]。近年では災害対応ロボット競技会 DARPA

Robotics Challenge (DRC) [34]をはじめとするヒューマノイドの災害環境対応における実証実現の場としても、車を運転する・ドアを回し操作で開ける・電動工具を操作する・バルブを回すなどのヒューマノイドによる環境操作実現が注目されている。車を運転するタスクのようなヒューマノイドが搭乗型機器の環境に拘束された状態下において、所望の動作生成を戦略し実現するに至るまでには多くの課題が存在する。その問題背景として、ヒューマノイドの四肢（双腕と双脚）および臀部などその他の部位の多点接触拘束が挙げられる。多点接触拘束下における動作生成・制御の研究は Harada ら [102] を筆頭に取り組みられており、解決となる提案手法が示されてきた [103][104]。

一方でこれらは外界にある接触対象として地面等の固定環境を対象としており、ヒューマノイドの接触する環境が操作などの作用に応じて動的に変化するような対象についてはあまり扱われていない。例えば、人間が日常生活でおこなっているような両手でハンドルを把持し操作するといった、運動拘束をもつハンドルに対して両手の接触拘束を保ちつつ動作を生成することなどが課題として挙げられる。以降では接触拘束下における動作生成法を背景に、ヒューマノイド身体と外界の静的/動的な環境との接触点連結により閉じた閉ループ系に着目し、その問題を分類・整理する。相山 [105] が提案した単マニピュレータにおける剛体閉ループ機構の概念に基づき、四肢をもったヒューマノイド動作へと拡張した閉ループ系の接触拘束について議論する。

静的な閉ループ系 Kanehiro らによる狭隘空間でのヒューマノイドのクローリング歩容 [106] の動作例（図 3.1 -(I)）では、双腕双脚とも床面の固定環境系に接地しておりヒューマノイド身体と床面の系で閉じた閉ループ系を示している。ここではヒューマノイド外界の環境が静的な固定環境のみで構成されている。本論文ではこのような系の接触状態を“静的な閉ループ系”と定義する。

準静的な閉ループ系 先述の DARPA Robotics Challenge (DRC) の車運転タスクの例 [37] では、図 3.1 -(II) のように片腕はハンドル操作のために作用し、もう片方の腕は固定環境となる助手席の座面に接地した方法がとられている。この場合、片腕は動的な環境であるハンドルへの作用素、もう片腕は固定環境接地の静的な非作用素となりハンドルの動的な反作用の影響を受けない。このような系を“準静的な閉ループ系”と定義する。

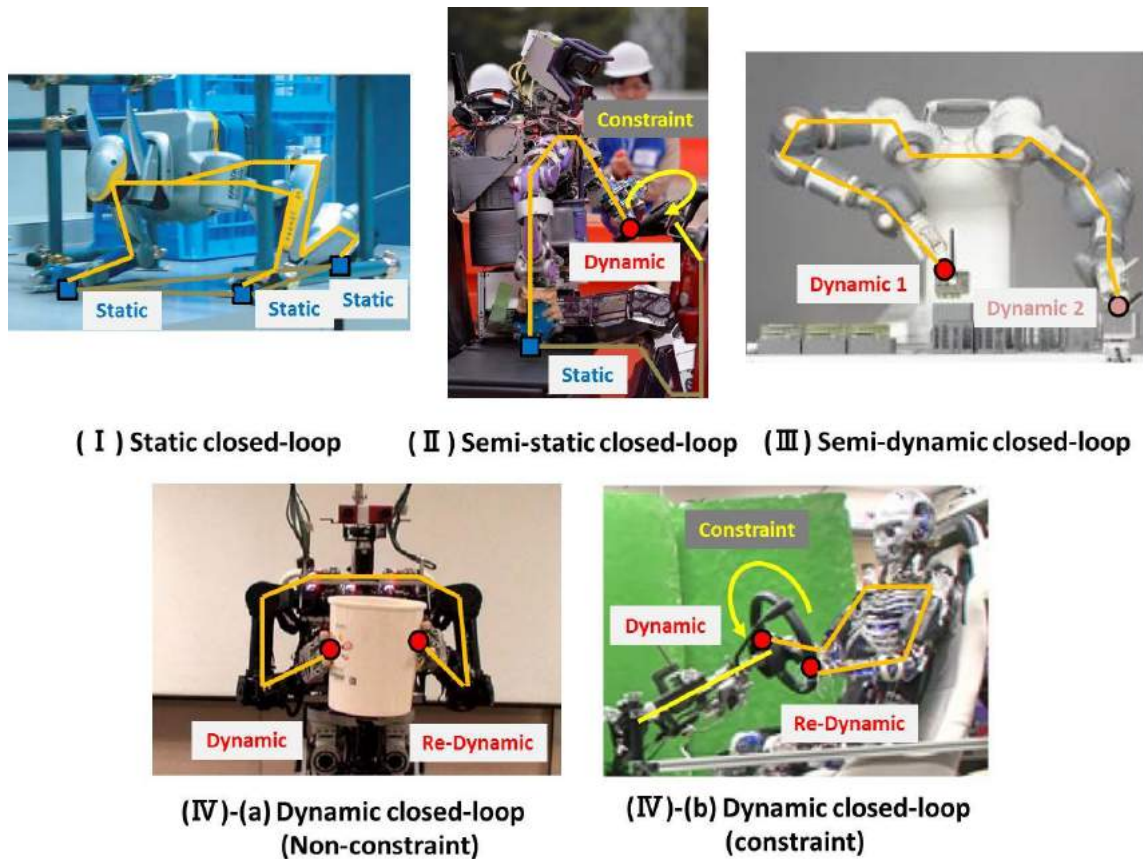


図 3.1: ヒューマノイドと接触対象における閉ループ系の分類

ヒューマノイド身体と接触対象から構成される閉ループ系の構造をそれぞれ, (I) 静的な閉ループ系 [106], (II) 準静的な閉ループ系 [37], (III) 準動的な閉ループ系 [107], (IV)-(a) 動的かつ接触対象非拘束な閉ループ系 [108], (IV)-(b) 動的かつ接触対象拘束な閉ループ系 [109] に分類する. (IV)-(a) および (IV)-(b) はまとめて“動的な閉ループ系”とも呼ぶ.

準動的な閉ループ系 閉ループ系において複数の動的な接触点を有するが, [107] (図 3.1-(III)) のように接触対象の運動が独立であるため, 系において反作用の影響を生じない. このときの系を“準動的な閉ループ系”と定義する.

動的な閉ループ系 最後に, 双腕によるハンドルの接触把持・操作 (図 3.1-(IV)-(a), 図 3.1-(IV)-(b)) に挙げられるような, 片方の作用素が動的な外部環境 (ここではハンドル) を介し, もう一方に反作用として影響を与える閉ループ系を“動的な閉ループ系”と定義する.

また, 動的な閉ループ系は作用・反作用の媒介層に位置する接触対象の運動の拘束条件についても考慮する必要がある. 双腕持ち上げ動作 [108] (図 3.1-(IV)-(a)) のように接触対象

物の運動が自由に記述できる場合を“動的かつ接触対象非拘束な閉ループ系”，ハンドルの回転運動などの運動制約をもった接触対象を操作する場合 [109] (図 3.1 -(IV)-(b)) を“動的かつ接触対象拘束な閉ループ系”とよぶ。

3.2.2 搭乗型機器の操作における閉ループ系の構造

前項で定義したヒューマノイドと接触対象から構成される閉ループ系構造の分類を基に，接触対象として本研究で注目する搭乗型機器を考える．ここでは多点接触となる搭乗型機器の例として，ヒューマノイドにおける三輪車操作の接触状態について閉ループ系の概念より観察していく．図 3.2 のように双腕双脚の四肢を有したヒューマノイドを三輪車に搭乗させた幾何的なモデルを考える．特徴となる接触点として，両手-両ハンドルグリップ，両足-両ペダル，臀部-サドルに近似した5点を与える．

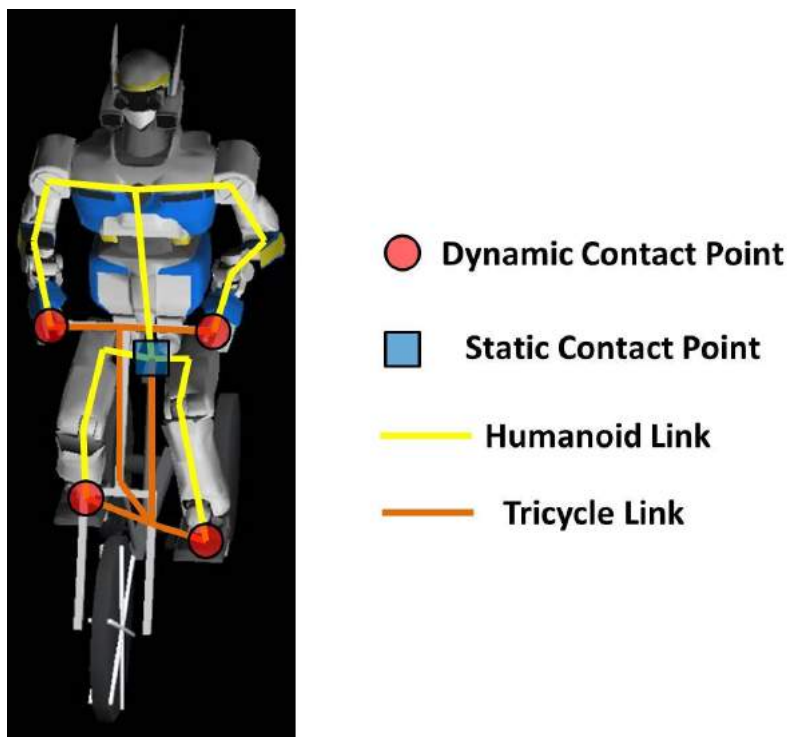


図 3.2: Closed-loop Contact Model For Humanoid HRP2-JSK [110] and Tricycle

ヒューマノイドと接触対象の間における閉ループ系の組み合わせの総数 S_{close} は，接触点の数を N_c としたとき，そのうちの任意の2点を選べば閉ループ系の構造が組まれることか

ら，組み合わせ総数 S_{close} は以下で算出される．

$$S_{\text{close}} = N_c C_2 = \frac{1}{2} N_c (N_c - 1) \quad (3.1)$$

したがって，ヒューマノイド身体と三輪車車体で構成される閉ループ系の組の総数は，全部で $S_{\text{close}} = 5C_2 = 10$ 種類の閉ループ系が存在していることがわかる．三輪車の車体を静的な固定環境，回転運動の拘束をもつハンドルおよびクランクを動的な接触対象としたとき，各閉ループ系の分類は図 3.3 の通りである．

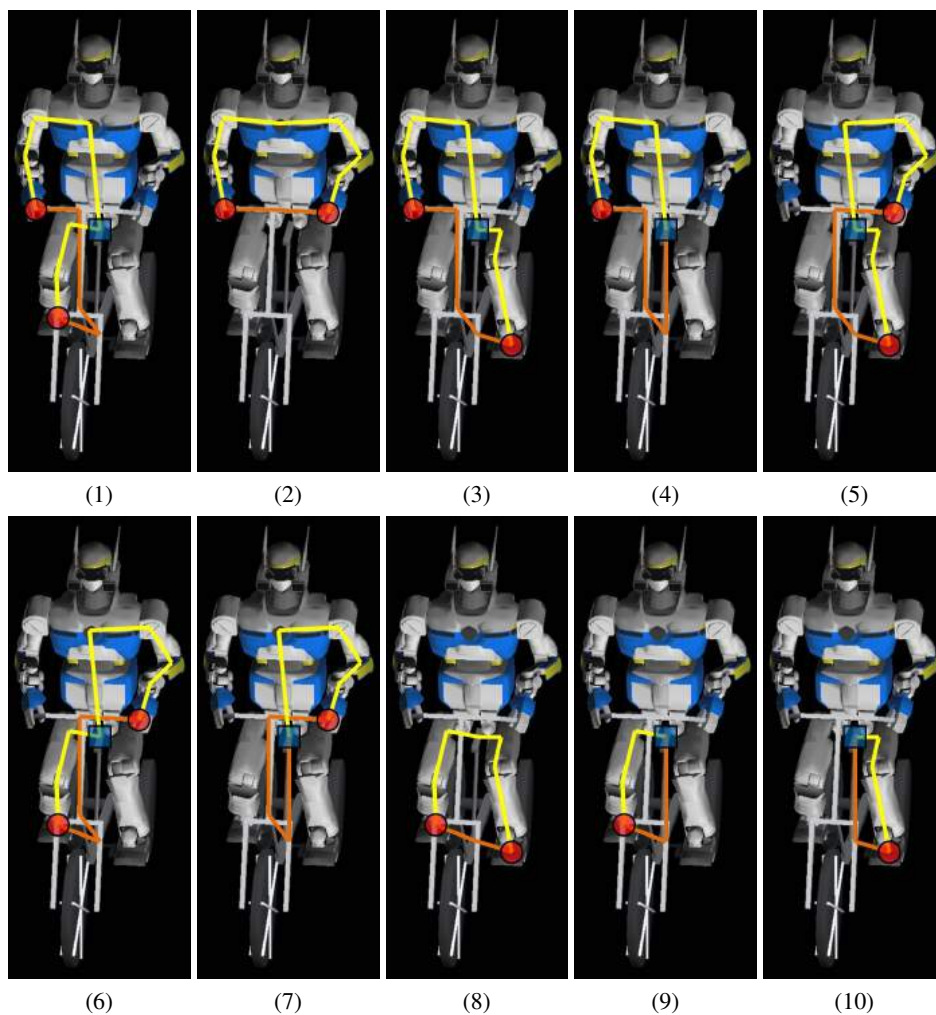


図 3.3: Classification of Closed-loop Systems for Humanoid and Tricycle

このうち，三輪車の操作に支配的なハンドルの把持接触点，クランクペダルの接触点のみから構成される閉ループ系は，準動的な閉ループ系と動的な閉ループ系の全 6 種類（図 3.3 -

(1),(2),(3),(5),(6),(8))である。特に、準動的な閉ループ系である4種類(図3.3-(1),(3),(5),(6))は閉ループ経路において臀部-サドルの経路接触点をもつため、図3.4に示すように経路点における閉ループ系の分割が可能となる。分割化された閉ループ系はすべて準静的な閉ループ系に分類される。

以上より、搭乗型機器として三輪車の操作で課題となる接触拘束は動的な閉ループ系の2種類(図3.3-(2),(8))である。これらは軸中心の回転運動の拘束をもった三輪車のハンドルおよびクランクに対しての、ヒューマノイドの双臂と双脚の操り(“動的かつ接触対象拘束な閉ループ系”での動作戦略)が要求される。

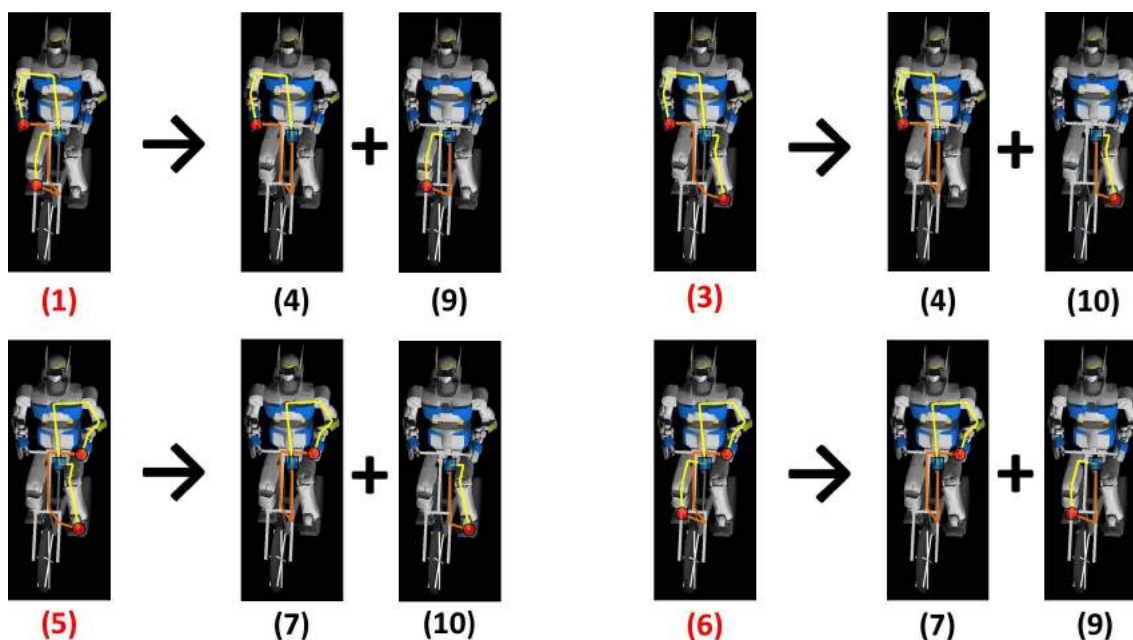


図 3.4: Classification Semi-dynamic Closed-loop (1),(3),(5),(6) to Semi-static Closed-loop (4),(7),(9),(10)

3.3 複数のマニピュレータを協調させた搭乗型機器の操作行動戦略

前節で述べた閉ループ系の接触拘束に関する議論を背景とし、ヒューマノイドがもつ複数のマニピュレータ(双臂と双脚)を協調させた搭乗型機器の操作のための行動戦略について本節で論じる。双臂・双脚の協調により操作される搭乗型機器の対象として本研究では三輪

車に注目し、三輪車のハンドルに対するステアリング操作行動とクランクに対するペダリング操作行動についてその特色を整理する。これらの議論を踏まえた上で、第2章で論じた探索指標の考えを用いたアプローチを示す。

3.3.1 三輪車の操作における課題

一般に三輪車を操作するという事は、ハンドルを切りクランクに回転運動を与えて漕ぎ進めることである。そのため、三輪車単体における操作の自由度は、独立したハンドルの回転およびクランクの回転の全2自由度と捉えることができる。

ヒューマノイドや人体模倣の対象となる人間が三輪車を操縦する場合、右手-右ハンドルグリップ、左手-左ハンドルグリップ、右足-右クランクペダル、左足-左クランクペダルの合計4入力を与えることができる。このような自由度数 N_f に比して制御入力数 N_u が多いシステム ($N_u > N_f$) は冗長システムとよばれる [111][112]。冗長システムのなかでも複数のマニピュレータと操作対象物の構成に着目した協調マニピュレーションの研究はロボティクス分野において盛んであり、吉田ら [113] による多腕強調制御や、Jeffrey ら [114] の多指ハンド解析などが挙げられる。マニピュレータの数が2つの場合の協調マニピュレーション操作は、一般に双腕マニピュレーションタスクとよばれる。

これまでに双腕マニピュレーションタスクの制御課題は中野ら [115] (図 3.5 -(I)) を始めとし、多様な提案手法が研究されてきた [116][117][118]。これらは主に拘束のない単一物体のマニピュレーションが前提であったが、吉川ら [119][120] により、対象物が平面拘束を受ける場合の操りの協調ハイブリッド制御系 (図 3.5 -(II)) が提案された。一方で、操作対象に拘束される運動に関して拡張を進めた研究は少ない。小菅ら [121] により拘束のある簡易的な組み付け作業を可能とする協調制御系が組み立てられているものの、日常で目にするような操作対象の回転操作等に対して、双腕マニピュレーションのアプローチにより実現した研究は数少ない。

本研究で対象とする三輪車のハンドルのステアリング操作やクランクのペダリング操作もヒューマノイドの双腕と双脚がなす双腕マニピュレーションタスクの一種である。これらの操作は一連の双腕マニピュレーションタスクの例においても、操作の対象物が運動拘束をもつ条件付きである。さらに、この運動拘束はハンドルやクランクの回転運動として表現され、先に述べた吉川ら [120] による平面拘束を受ける対象物の操りからの運動拘束の拡張・発展が求められる。特に回転運動として記述される拘束下での双腕マニピュレーションタス

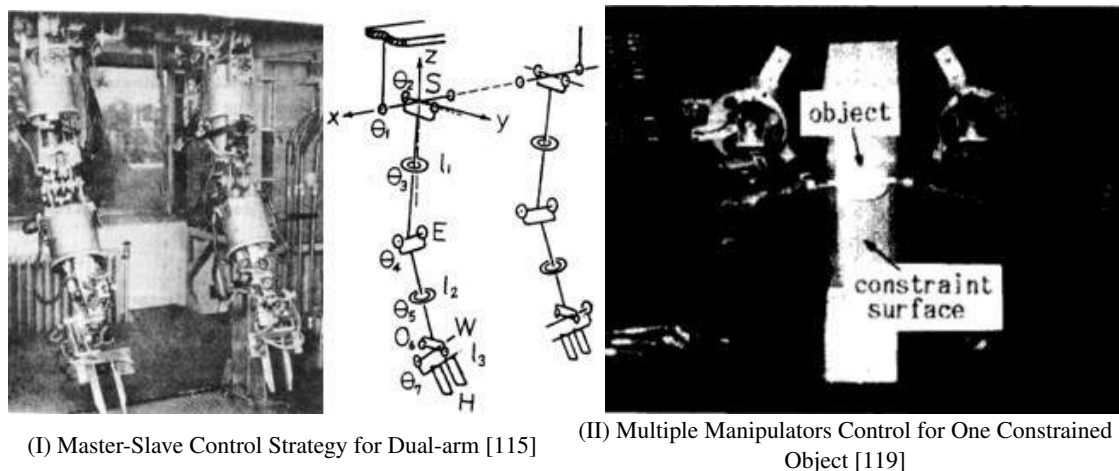


図 3.5: Previous Research Works about Multiple Arms Coordinated Manipulation

クの実現においては、片方のマニピュレータの挙動がもう一方のマニピュレータの動作生成に大きく影響を与える問題がある。

以上を整理すると、三輪車の操作においては次に示す複合的な課題を解く必要が挙げられる。

1. 冗長システム ($N_u > N_f$) [111][112]
2. 協調マニピュレーションタスク [113][114]
3. 双腕マニピュレーションタスク（操作対象に支配的なマニピュレータの数が2つの場合における協調マニピュレーションタスク） [115][116][117][118]
4. 操作対象に拘束された運動が存在 [119][120][121]
5. 操作対象に回転運動の拘束が存在（操作対象に拘束された運動の拡張）
⇒ 一方のマニピュレータの挙動ともう一方のマニピュレータの挙動が大きく干渉

上記のような束縛された条件は、3.2節で定義・分類した“動的かつ接触対象拘束な閉ループ系”として捉えることが可能であり、この問題に対する新たな制御戦略および具体的な方策について検討し提案していく必要があると言えよう。

3.3.2 ヒューマノイドの双腕によるハンドルのステアリング操作

前項の課題を踏まえた制御方策の構築の準備として、具体的な操作行動戦略を観察する。ヒューマノイドの双腕を用いた三輪車のハンドルのステアリング操作に対するリンク機構の構成図を図3.6-(I)に示す。リンク機構はヒューマノイドの双腕リンクおよび胴体リンクと、三輪車のハンドルのリンクから構成された閉リンク構造であり、ハンドルシャフトが回転関節に相当する。ここでの腰リンクは、固定された環境を仮定したサドルと結合関係にあるとする。両手はハンドルグリップを把持した状態であり、閉ループ系の拘束下において任意の方向に位置および力が発揮される。

構造に示す通り、ハンドルには回転運動の拘束が存在するため、双腕のエンドエフェクタ同士は互いに大きく影響し合うことが観察できる。特に両手先のエンドエフェクタの位置制御等を想定した場合、計画した軌道の厳密な作成が困難であり、実環境におけるハンドルの回転運動を精密にモデル化することは難しい。なおかつ搭乗する三輪車が異なった場合や、三輪車に対するヒューマノイドの搭乗位置・姿勢が変わった場合には逐一、計画する軌道を吟味する必要がある。煩雑性も懸念事項となる。ヒューマノイドの手先に力制御を適用した場合にも同様のことが言える。以上の観点から、ヒューマノイドによるハンドルのステアリング操作においては双腕の各機能を

- 操作に必要なとされる入力軌道を計画し実行する作用要素
- 操作対象（ハンドル）の拘束された運動に対して受動的に軌道を獲得していく反作用要素

に切り分けた操作行動戦略をとることがポイントとなる。

3.3.3 ヒューマノイドの双脚によるクランクのペダリング操作

足裏とペダルの接触抗力の満足

ヒューマノイドの双脚がなす三輪車のクランクのペダリング操作に注目する。図3.6-(II)がペダリング操作におけるリンク機構の構成であり、ヒューマノイドの双脚リンクと臀部リンク（腰リンク）、および三輪車のペダル・クランクのリンクで閉じた系となる。ハンドルのステアリング操作の場合と同様に、臀部リンク（腰リンク）は三輪車の車体の環境に固定

されているものとする。両足裏はビンディングなどは設けずにペダルと接触しており、接触状態の離脱が考えられる。

ペダリング操作に関しても、ペダル・クランクを介して両足先のエンドエフェクタ同士が干渉し合う構造となっている。先に述べたステアリング操作と異なり、エンドエフェクタが三輪車のペダルと把持状態ではなく、干渉を回避するような接触状態の離脱が可能であると言えるが、離脱による操作性のズレの伝搬やヒューマノイドの搭乗姿勢の不安定性へと影響が及ぼされる。このことから、ヒューマノイドの足裏と三輪車のペダルの間で発生する接触抗力を十分に保つことが重要なポイントとなる。

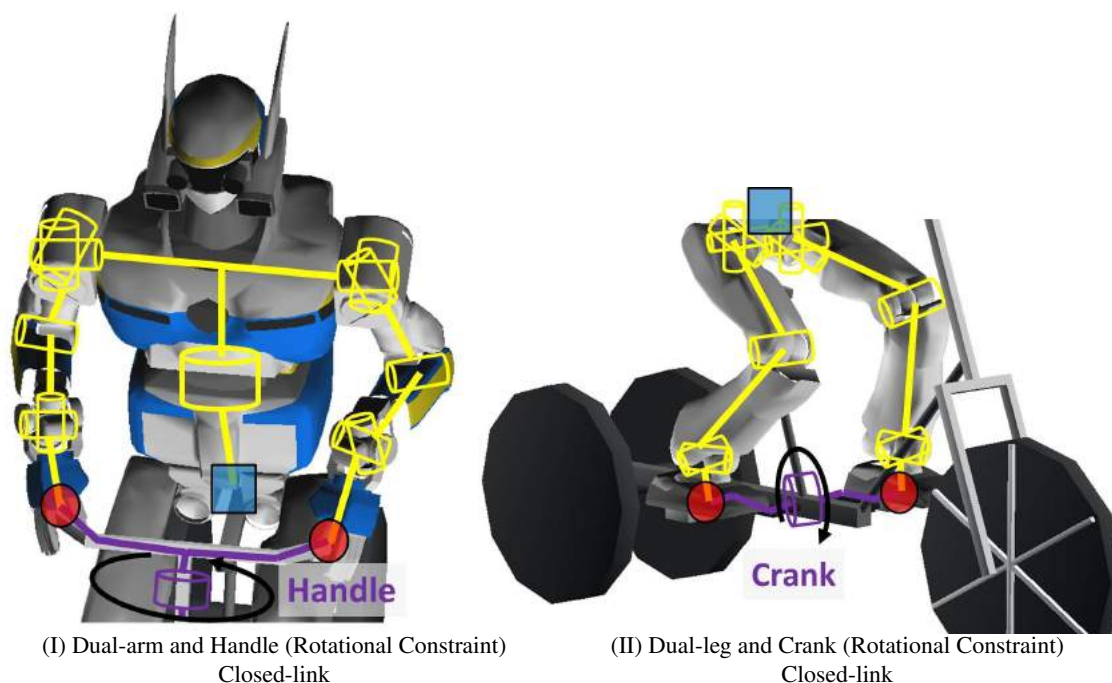


図 3.6: Closed-link System Composed of Dual-arm and Handle, Dual-leg and Crank

ペダリング操作における機能の分担と切替行動

本研究では、先行研究 [122][123][124] のようにビンディングペダルの適用は考えず、一般的な日常生活におけるシティサイクルタイプの三輪車を想定する。クランクのペダリング操作は踏み込みによる操作が支配的であると見做し、引き脚を利用した操作行動はおこなわないものとする。したがって、ハンドルのステアリング操作の項で述べた2つのマニピュレータの

作用要素と反作用要素は、双脚がなす運動においては踏み込みのための脚とそれに対応した接触抗力を維持するための脚へと一意に機能の割り当てを決定することができる。

Gfohlerら [125] によるサイクリング運動時の FES（機能的電気刺激: Functional Electrical Stimulation）の動力学シミュレーションでは、図 3.7 に示すように人間の下肢筋肉のなかでも運動能力として貢献度の高いハムストリングス（HAM）の能動期（クランクに対しての駆動トルクが高まる ε の区間）として、図 3.7-(a) のようなパターンを与えている。本研究においてもこの区間パターンを参考に、踏み込みの脚の能動期を $20[\text{deg}] \leq \varepsilon \leq 200[\text{deg}]$ として与える。したがって、マニピュレータの作用要素と反作用要素の機能を切り替えるクランク角度を $\varepsilon = 20[\text{deg}]$ とした切替行動を戦略する。

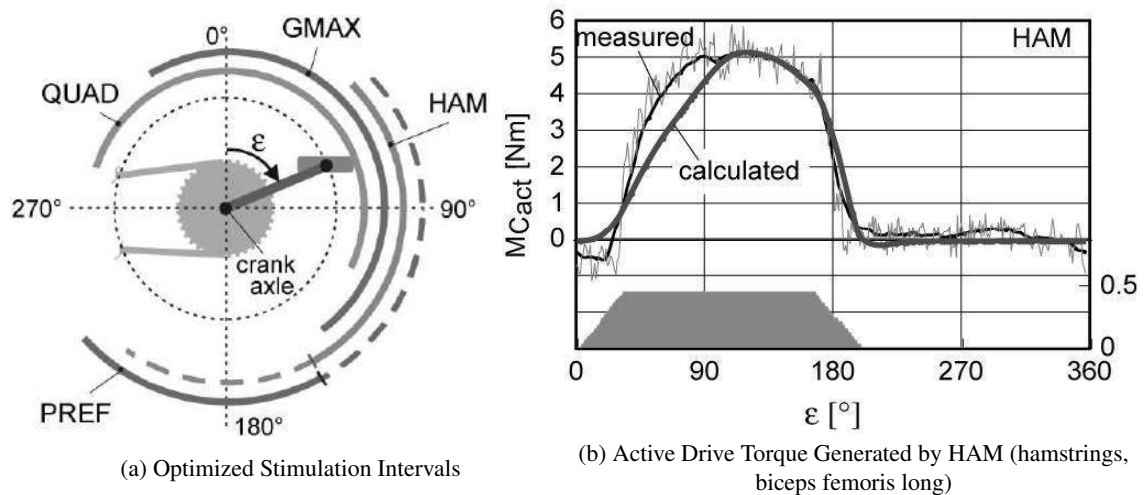


図 3.7: Dynamic Simulation of FES-Cycling [125]

以上から、ヒューマノイドの双脚によりペダリング操作を実現するためのポイントをまとめると以下のようなになる。

- 離脱を防止するための足裏とペダルの間における接触抗力の十分な維持
- 踏み込みを行う脚を作用要素，接触抗力を維持する脚を反作用要素とした機能分担
- 作用要素と反作用要素の切替方策

3.3.4 ステアリング操作とペダリング操作における共通行動と特徴

共通行動

双腕によるハンドルのステアリング操作と、双脚によるクランクのペダリング操作の共通行動として、作用要素と反作用要素の制御機能の切り分けが根幹をなしている。これらの機能分担は、“動的かつ接触対象拘束な閉ループ系”を呈するステアリング操作とペダリング操作に対しての双腕と双脚による協調マニピュレーションの解となり得る。特に、先行研究 [126] におけるヒューマノイド脚部のみによるペダリング操作を、操舵のためのステアリング操作に対しても適用可能な共通の制御方策として拡張することは非常に大きな有用性が得られる。

このような行動戦略は三輪車のハンドルとクランクに対するステアリング操作・ペダリング操作に留まらず、その他の回転運動の拘束を有した対象物についても同様に適用可能であるといえる。例えば、災害対応ロボット競技会 DARPA Robotics Challenge (DRC) [34] のタスク課題としても与えられているヒューマノイドによるバルブの回し操作に関しては、先行研究 [127][128] のように片腕のみで動作生成を実行する手法が提案されているが、人間と同じように両腕を用いて回転剛性の大きいバルブを回していくことが課題となる。円形バルブは軸中心の回転操作が可能な機構であるため、図 3.6 で示したようなハンドルやクランクの操作と同様の操作行動戦略のアプローチをとることができる。以下に共通の行動戦略として適用が可能な回転操作タスクの行動例を示す。

- 円形バルブの回し操作
- 車のハンドル操舵
- 運搬用台車の旋回操作

上記に加えて、異なる運動拘束をもつ操作対象としてヒューマノイドによる箒掛け動作 [129] など幅広い行動に対して汎用性が挙げられる。

接触拘束の状態の特徴

ハンドルのステアリング操作とクランクのペダリング操作はともに“動的かつ接触対象拘束な閉ループ系”に分類されるが、双腕と双脚のエンドエフェクタに相当する手先・足先における接触の状態が異なる。接触状態の分類 [130] に基づいて、両手先-ハンドル、両足先-

クランクの接触状態を考えると、それぞれ握り接触、線接触といえる。ここで、ペダルへの接触姿勢は受動関節よりフリーであるとする。

握り接触となる手先とハンドルグリップの接触では、6自由度の拘束が約束されており、各自由度方向に対して接触対象に作用力を与えることができる。

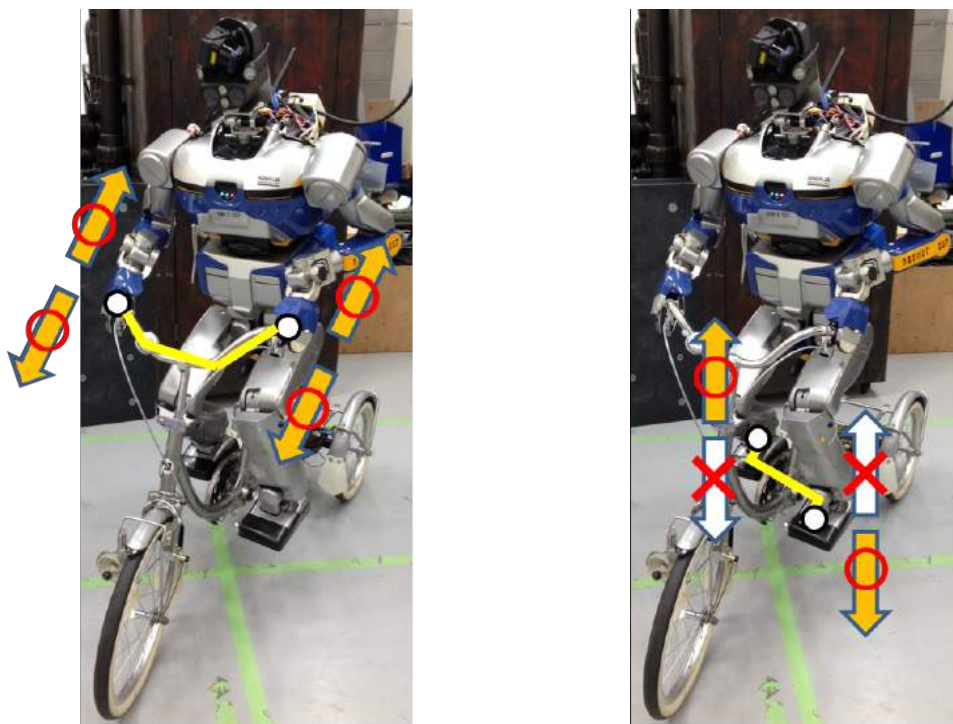
線接触である足先とクランクでは、接触対象となるペダルシャフト面のローカル座標系において x 軸まわりの回転 θ_x のみの拘束が約束されている。足裏とペダル面の静止摩擦による接触拘束はここでは考慮しないものとする。位置拘束は z の片方向 ($-z$ 方向) へのみ存在し、このような拘束を“ユニラテラル拘束”とよぶ。反対に正負両方向に対する拘束を“バイラテラル拘束”とよぶものとする（例えば、 θ_x の拘束はバイラテラル拘束に分類される）。ユニラテラル拘束の場合、一方向に対してのみ接触対象への作用が可能である。

手先とハンドルグリップの接触拘束状態は、握り接触かつすべての拘束がバイラテラル拘束である。一方で足先とクランクの接触拘束の状態は、線接触かつ1つのユニラテラル拘束をもつ。そのため、双腕によるハンドルのステアリング操作では、各腕に対する作用要素・反作用要素の機能適用を任意に決めることができ、双脚によるクランクのペダリング操作では、ユニラテラル拘束の存在から、作用要素と反作用要素を踏み込みの脚ともう一方の脚のように一意に決定しなければならない（図 3.8）。

以上がステアリング操作とペダリング操作の大きな異なる特徴である。表 3.1 にこれまでに紹介した、双腕によるハンドルのステアリング操作と双脚によるクランクのペダリング操作の差異をまとめる。

表 3.1: Difference between Steering and Pedaling

	ステアリング操作	ペダリング操作
接触で拘束される自由度	6-DoF ($x, y, z, \theta_x, \theta_y, \theta_z$)	2-DoF (z, θ_x)
拘束の種類	$x, y, z, \theta_x, \theta_y, \theta_z$ 全てにおいてバイラテラル拘束	z はユニラテラル拘束, θ_x はバイラテラル拘束
作用要素/反作用要素の機能分担	任意に決定（右腕/左腕どちらでも可）	一意に決定（踏み込みの脚/反対側の脚）
機能切替の必要性	不要（任意でハンドルの切替角を指定）	必要（クランクの切替角 $\varepsilon = 20[\text{deg}]$ ）



ステアリング操作における
バイラテラル拘束

ペダリング操作における
ユニラテラル拘束

図 3.8: Steering with Bilateral Constraint and Pedaling with Unilateral Constraint

3.3.5 探索指標の考えに基づいた操作行動戦略

これまでに述べたように三輪車の搭乗型機器操作の例においては、ヒューマノイドの複数のマニピュレータ（双腕・双脚）を協調させた閉ループ系での回転運動拘束をもつ対象（ハンドル・クランク）の操作行動が課題となる。この課題は3.2節で定義した“動的かつ接触対象拘束な閉ループ系”として分類され、接触対象が回転運動の拘束をもつ場合、双方のマニピュレータ同士の動作生成の軌道が互いに干渉し合う問題が存在することについても述べた。この問題を解くにあたって、我々は第2章で詳説したオンライン探索指標を主軸とする探索指標の考え方を取り入れ、上記の協調マニピュレーションの課題解決に取り組む。

操作に支配的な2つのマニピュレータに対して、図3.9に示すような「能動機能」と「受動機能」の2種類の制御機能の考えを導入する。「能動機能」の実現にはオフライン探索指標を利用し、「受動機能」の実現に関してはオンライン探索指標によるアプローチを提案す

る。各機能の定義および探索指標の概念を用いた実現法については3.4節で詳説するものとして、以下に簡単な機能概要を示す。

能動機能 操作対象物体に所望の操作入力（軌道）を与えるマニピュレータの制御機能
⇒ “オフライン探索指標” を利用することで実現

受動機能 操作対象物体の運動に応じてオンラインで軌道を獲得するマニピュレータの制御機能
⇒ “オンライン探索指標” を利用することで実現

これらの制御機能は、ヒューマノイド双腕-三輪車ハンドル、ヒューマノイド双脚-三輪車クランクの2つの組で共通に適用可能な機能である。上記2つの機能を用いることで、脚部の協調動作のみに注目したヒューマノイドのペダリング操作の研究 [126] から拡張された、搭乗型機器の進行と操舵の2つの基本要素を満たす運転操作行動が可能となる。

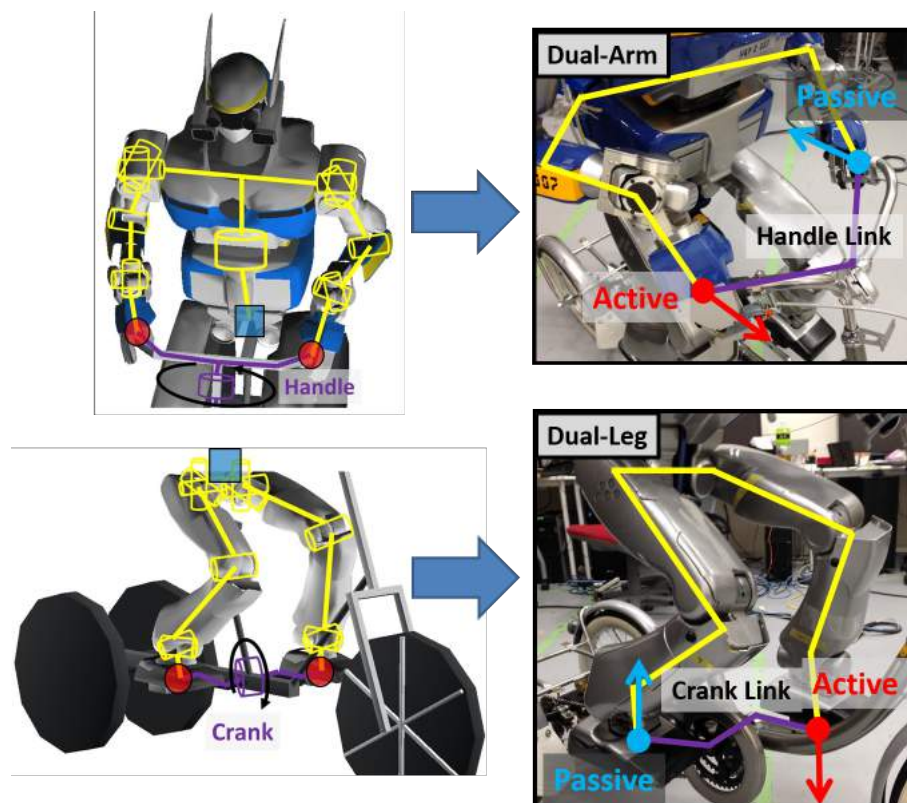


図 3.9: Active/Passive Functions for Steering and Pedaling Tasks by Dual-Manipulation [52]

3.4 探索指標を利用した能動機能と受動機能のマニピュレータ軌道における制御法

3.3.5 項で述べた探索指標を用いた操作行動戦略に関して具体的な制御法を本節で提案する。双腕や双脚のような協調する2つのマニピュレータに対して，“能動機能”と“受動機能”の新たな制御機能を定義・提案する。3.3.5 項で示した機能概要のように，能動機能の制御の実現は“オフライン探索指標”を利用し，受動機能の制御の実現では“オンライン探索指標”を利用した制御戦略をとる。これら2つの探索指標に基づくマニピュレータの能動機能と受動機能により，搭乗型機器の操作に対する軌道計画と軌道修正を実現させる。

目的となるハンドルのステアリング操作やクランクのペダリング操作における動作は，“動的かつ接触対象拘束な閉ループ系”の拘束制約をもつことが3.2節で議論された。さらに，その動作生成の方策として双方のマニピュレータに関して，必要な操作入力を対象に与える作用要素と，対象の運動に応じて受動的に追従し生成軌道を獲得する反作用要素に分解する考えの重要性が3.3節で述べられた。本節では，これらのマニピュレータがもつ機能要素の動作制御法について，第2章で詳しく論じた探索指標によるアプローチに基づいて提案する。

3.4.1 マニピュレータの能動機能と受動機能の定義

ヒューマノイドの双腕や双脚を一般化した2つのマニピュレータによる協調マニピュレーションのモデルを図3.10に示す。図3.10の協調するマニピュレータがもつ作用要素と反作用要素が呈する動作機能を，本論文では“能動機能”と“受動機能”として定義する。マニピュレータにおける能動機能と受動機能の2つの概念を一般化して定義すると以下のように表現される。

能動機能の一般化定義 ロボットマニピュレータと操作対象物体は接触状態にあるとする。操作対象物体の所望運動 $\Sigma(t)$ から導かれるロボットマニピュレータと操作対象の接触(把持)集合 C_A の運動 $\Sigma_{C_A}(t)$ を求める。運動 $\Sigma_{C_A}(t)$ に対して，ロボットマニピュレータ上の C_A が生成すべき位置・姿勢または力・モーメントの入力軌道 $\mathbf{X}_{\text{active}}^{\text{ref}}(t)$ を考える。このとき，軌道 $\mathbf{X}_{\text{active}}^{\text{ref}}(t)$ への追従を満足するロボットマニピュレータの入力列 \mathbf{u} を求め，入力列 \mathbf{u} の再生が可能であるとき，このロボットマニピュレータは能動機能を有するという。

受動機能の一般化定義 今ロボットマニピュレータにはインピーダンス制御 [131][132][49][50] が適用されているものとする。操作対象物体が $\Sigma(t)$ の運動 (拘束を含む) をしているとき、ロボットマニピュレータと対象物体の接触集合 C_P は $\Sigma_{C_P}(t)$ の運動を形成する。このとき、接触集合 C_P においてマニピュレータは操作対象物体より外力・外モーメント $W_{C_P} = (F_{C_P} M_{C_P})^T$ を受ける。接触集合 C_P がロボットマニピュレータのエンドエフェクタ集合 E に存在 ($C_P \in E$) する場合、インピーダンス制御が有効であり、外的要因 W_{C_P} によりロボットマニピュレータの位置姿勢は $\xi(t)$ に遷移する。ここで t は現時刻であり、次の時刻 $t + \Delta t$ における目標位置姿勢 $\xi^{ref}(t + \Delta t)$ が $\xi^{ref}(t + \Delta t) = \xi(t)$ を満たすように追従可能であるとき、このロボットマニピュレータは受動機能を有するという。

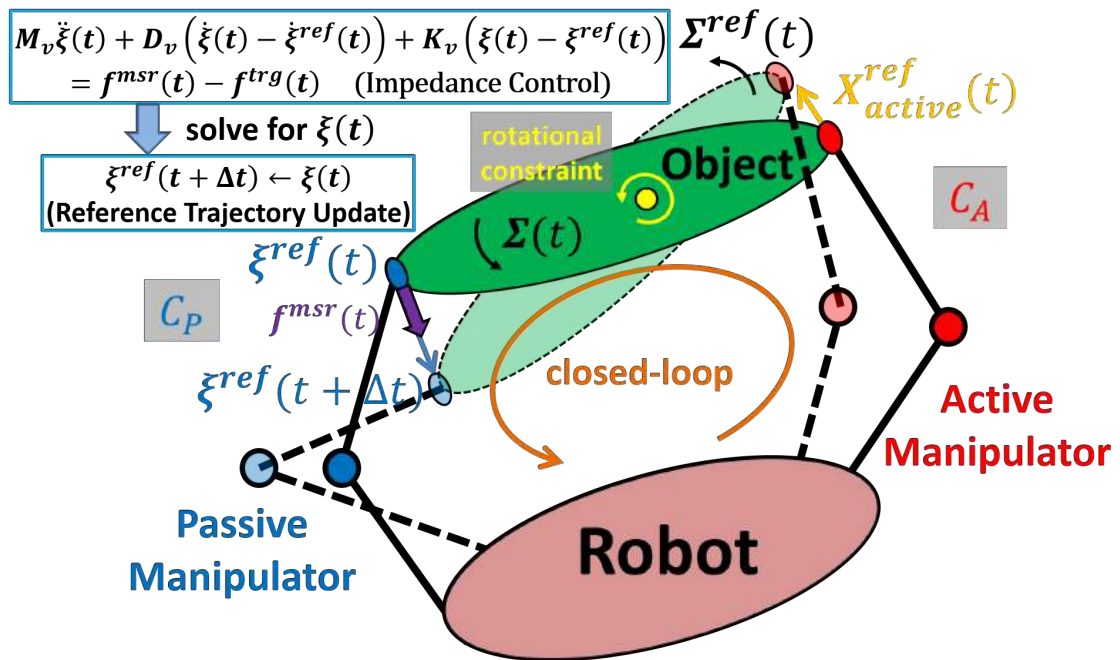


図 3.10: Manipulation Control Model for Active/Passive Manipulators

能動機能は操作対象の目標運動からオフラインに生成する軌道計画法を採用しており、受動機能は操作対象の実環境における運動に追従するようなオンラインの軌道獲得法を適用している。以降では、オフライン探索指標の考えに基づいた能動機能、オンライン探索指標の考えに基づいた受動機能の各々の機能の制御法を具体的に述べていく。

3.4.2 オフライン探索指標に基づく能動機能の制御則

操作対象物体の運動と拘束

能動機能の一般化定義に従って，操作対象物体の所望となる運動を考える．三輪車を例とした操作の自由度は前述の通り，ハンドルの回転運動およびクランクの回転運動の2自由度である．図 3.11 のようなヒューマノイドが搭乗する三輪車の形状および運動学モデルを与える．

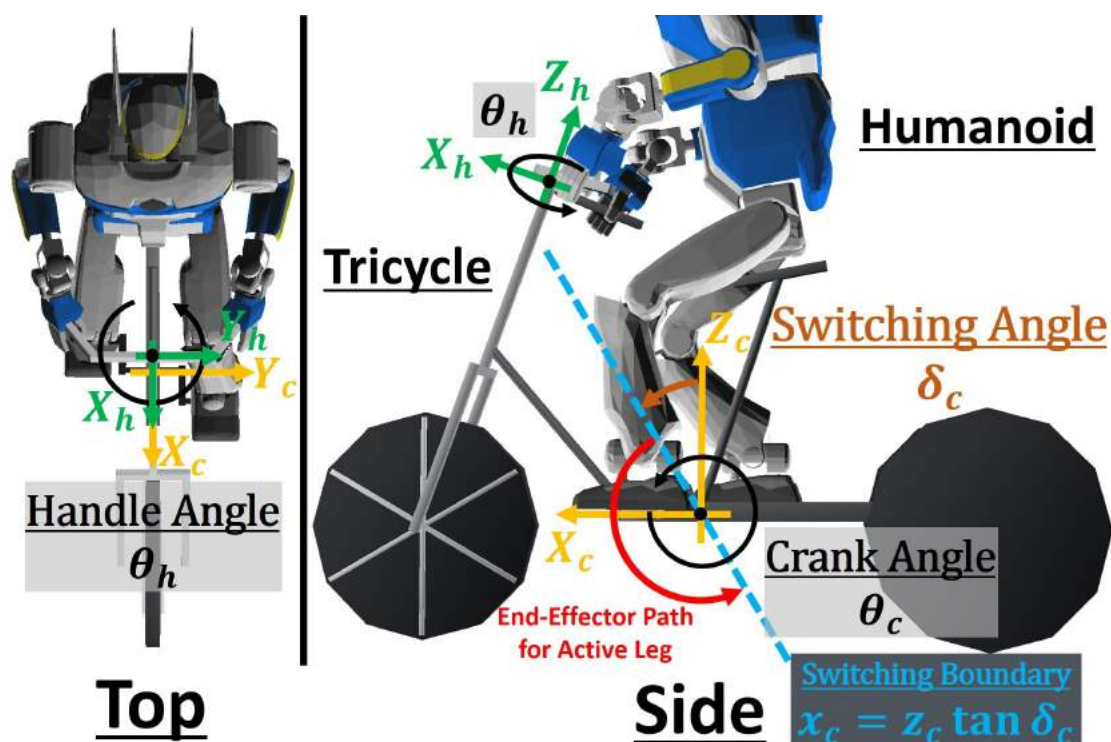


図 3.11: Kinematics Model for Tricycle with Humanoid HRP2-JSK [110] [53]

Manipulation parameters for tricycle are handle angle θ_h and crank angle θ_c . Switching crank angle for active/passive phase is defined as δ_c .

一般ハンドル角を θ_h [rad]，一般クランク角を θ_c [rad] とするとき，能動機能のマニピュレータが通過する θ_h, θ_c の取りうる値の範囲は次のように与えられる．ただし，図中での三輪車の姿勢を $\theta_h = 0$ [rad]， $\theta_c = 0$ [rad] とし，ハンドル・クランクの座標系も図 3.11 にしたがう．

$$-\Theta_H \leq \theta_h \leq \Theta_H \quad (3.2)$$

$$-\infty \leq \theta_c \leq \infty \quad (3.3)$$

ここで、 Θ_H はハンドルの最大操舵角度であり、三輪車固有の物理的な拘束角度 Θ_{H0} に加えて、搭乗するヒューマノイドが自己身体干渉しないための制約角度 $\Theta_H(\mathbf{K}_{\text{humanoid}}, \mathbf{q}_{\text{humanoid}})$ の存在も考慮する必要がある ($\mathbf{K}_{\text{humanoid}}$: ヒューマノイドの Kinematics パラメータ群, $\mathbf{q}_{\text{humanoid}}$: ヒューマノイドの姿勢列). したがって、 Θ_H は以下を満たすパラメータとなる.

$$\Theta_H = \min(\Theta_{H0}, \Theta_H(\mathbf{K}_{\text{humanoid}}, \mathbf{q}_{\text{humanoid}})) \quad (3.4)$$

さらに、クランクの角度 θ_c は周期性より以下のように有界制約に書き表すことができる.

$$0 \leq \theta_c < 2\pi \quad (3.5)$$

以上より、操作対象物体であるハンドル・クランクの運動拘束に関する制約は式 3.2, 式 3.4, 式 3.5 を満たす回転運動であるといえる.

オフライン探索指標を利用した目標軌道の設計

拘束制約下における所望運動 $\Sigma(t)$ と、前項で示したハンドル・クランクの運動拘束および Kinematics パラメータ群の写像に基づき、ハンドルグリップの把持点とクランクペダルの接触点を通る軌道の集合 $\mathbf{X}_{\text{active}}^{\text{ref}}(t)$ が形成される. ここで、クランクのペダリング操作においては、クランクペダルの接触点の通過領域を (x_c, y_c, z_c) とし、Gfohler ら [125] によるサイクリング運動の区間パターンを参考に、踏み込み期の存在を考慮した能動機能の適用範囲を以下の制約条件として与える.

$$x_c \geq z_c \tan \delta_c \quad (3.6)$$

δ_c は文献 [125] におけるパラメータ ε に相当する z 軸からの偏角であり、本研究では能動期と受動期の切替角として $\delta_c = 20 \cdot \frac{\pi}{180} [\text{rad}]$ と設定する.

上記の制約をもとに、クランクのペダリング操作における目標軌道設計を行う。離散的な時刻 i におけるヒューマノイドの能動機能の脚の足先位置を $\mathbf{p}_i = (x_{pi} \ y_{pi} \ z_{pi})^T$ とし、図 3.12 のようにオフラインで設計した足先の軌道を描く。

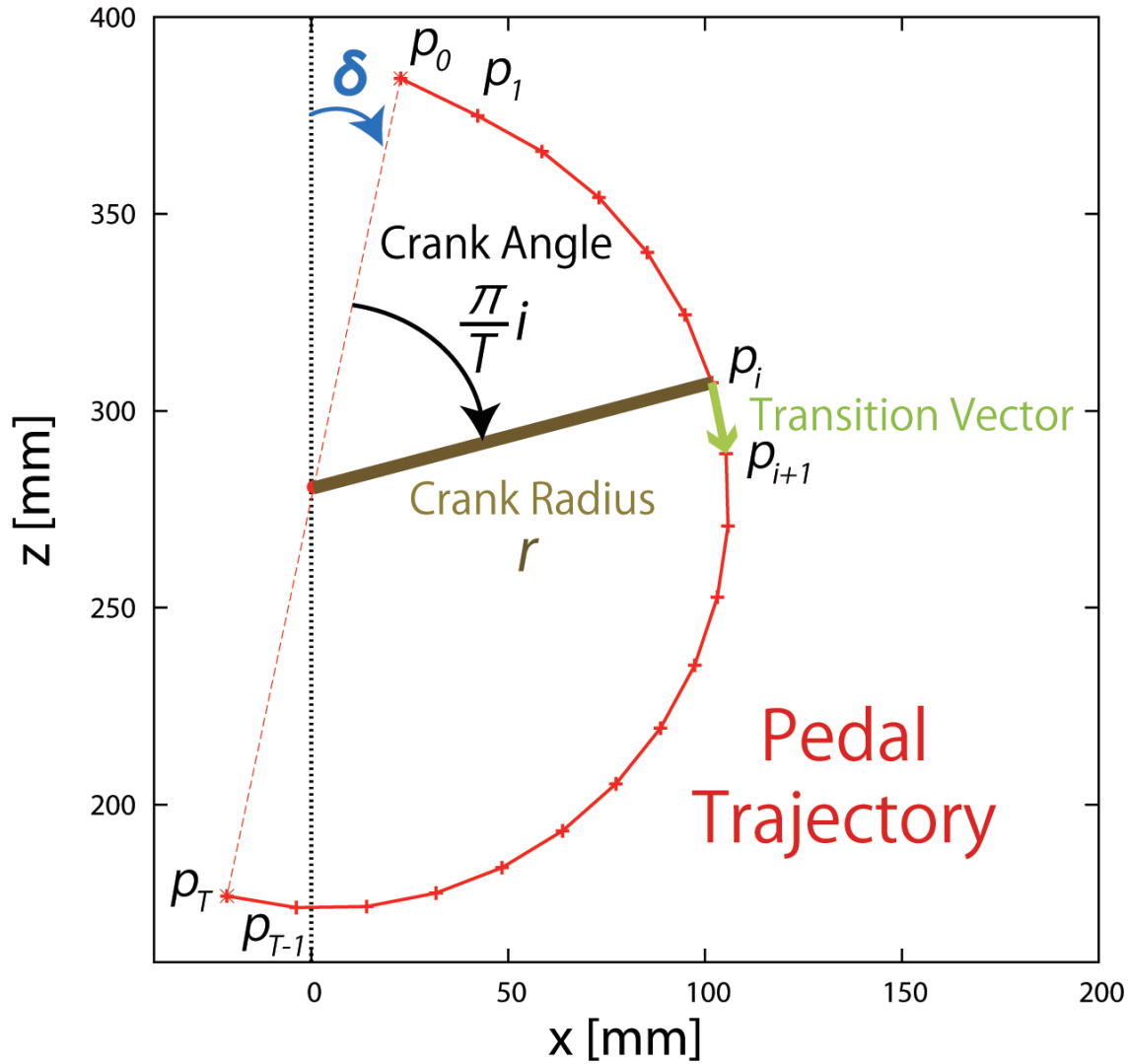


図 3.12: Offline Designed Trajectory for Foot End-Effector [126]

能動期の始点（ペダリング操作における上死点）の位置を \mathbf{p}_0 、能動期の区間 $\frac{\pi}{2} + \delta_c \leq \theta_c < \frac{3\pi}{2} + \delta_c$ における軌道のサンプリングの区間数を T 、クランク半径を r としたとき、 \mathbf{p}_i は以下

で記述できる.

$$\begin{aligned} \mathbf{p}_i &= \mathbf{p}_0 + r \left(\mathbf{R}_y \left(\frac{\pi}{T} i \right) - \mathbf{I}_3 \right) \mathbf{v}(\delta_c) \\ &, i = 0, 1, \dots, T-1, T \end{aligned} \quad (3.7)$$

ここで \mathbf{I}_3 は $\mathbb{R}^{3 \times 3}$ の単位行列を表し, $\mathbf{R}_y(\theta), \mathbf{v}(\delta_c)$ はそれぞれ式 (3.8) の行列, ベクトルで表現される. 行列 $\mathbf{R}_y(\theta)$ は y 軸 (クランク軸) まわりに θ だけ回転させる回転行列を表す.

$$\begin{aligned} \mathbf{R}_y(\theta) &= \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \\ , \mathbf{v}(\delta_c) &= \begin{pmatrix} \sin \delta_c & 0 & \cos \delta_c \end{pmatrix}^T \end{aligned} \quad (3.8)$$

離散時刻 i におけるクランクの角度を $\frac{\pi}{T}i$ として, 0 から π まで回転操作の軌道がオフラインで生成される. ロボットの足先位置はこの計画した軌道に対して追従するように位置制御が実行される.

設計した目標軌道の更新則をオフライン探索指標の考えを用いて導出する. 式 3.7 より現在の足先位置 \mathbf{p}_i と次ステップの目標の足先位置 \mathbf{p}_{i+1} の関係を導くと以下の更新式が得られる.

$$\mathbf{p}_{i+1} := \mathbf{p}_i - r \left(\mathbf{R}_y \left(\frac{\pi}{T} i \right) - \mathbf{R}_y \left(\frac{\pi}{T} (i+1) \right) \right) \mathbf{v}(\delta_c) \quad (3.9)$$

式 2.13 で定式化した“軌道とゲインの一般化更新則”に当てはめて考えると, 以下のような対応であることがわかる.

表 3.2: “軌道とゲインの一般化更新則” (式 2.13) における対応

オフライン探索指標	更新される軌道	一般化更新則の重み A
$\mathbf{R}_y \left(\frac{\pi}{T} i \right) - \mathbf{R}_y \left(\frac{\pi}{T} (i+1) \right)$	\mathbf{p}_i	$r, \mathbf{v}(\delta_c)$

$\mathbf{R}_y \left(\frac{\pi}{T} i \right)$ が現在状態, $\mathbf{R}_y \left(\frac{\pi}{T} (i+1) \right)$ が目標状態に相当する. $\frac{\pi}{T}i$ は現在時刻 i におけるクランク角度の状態を表しており, 探索変数に関連したパラメータである. このようにオフライン探索指標を用いることで所望とする目標軌道の設計が行える.

制御入力の決定

最後に、ヒューマノイド四肢のエンドエフェクタが軌道 $\mathbf{X}_{\text{active}}^{\text{ref}}(t)$ に追従するような入力列 \mathbf{u} を算出することで、能動機能を有するマニピュレータの制御入力決定される。ここでの入力列 \mathbf{u} は、ヒューマノイドの関節入力列であり、目標軌道 $\mathbf{X}_{\text{active}}^{\text{ref}}(t)$ で決まるエンドエフェクタの目標位置姿勢からの逆運動学により求めることができる。

入力列 \mathbf{u} が満たすべき制約は

1. 入力制限（関節角度のリミット、関節の定格トルク）
2. 目標軌道 $\mathbf{X}_{\text{active}}^{\text{ref}}(t)$ が可到達領域（Reachability Map [133]）内に存在
3. 軌道追従時において自己身体干渉のないハンドル操作が可能

であり、定式化すると次のように表せる。

$$\mathbf{u} \in U_0 \quad (3.10)$$

$$FK(\mathbf{u}) = \mathbf{X}_{\text{active}}^{\text{ref}}(t) \in \text{ReachabilityMap} \quad (3.11)$$

$$\forall t, \exists \mathbf{u} \text{ s.t. } \Theta_H(\mathbf{K}_{\text{humanoid}}, \mathbf{q}_{\text{humanoid}}) > 0 \quad (3.12)$$

それぞれ U_0 : 入力制限集合, $FK(\cdot)$: 順運動学写像, ReachabilityMap : 可到達領域 (Reachability Map) 集合を表す。

3.4.3 オンライン探索指標に基づく受動機能の制御則

制御の適用可能性

受動機能の特色として、操作対象物体の運動学モデル・動力学モデルの情報を一切与えずに、実際の運動（実環境での運動）に応じて軌道を獲得・生成していくオンライン性が挙げられる。図 3.13 のように操作対象物体の動きはブラックボックスであるが、運動する対象物体から受ける外的要因（力・モーメント） $\mathbf{W}_{\text{obj}} = (\mathbf{F}_{\text{obj}} \ \mathbf{M}_{\text{obj}})^T$ が明示的であり、かつ \mathbf{W}_{obj} がマニピュレータのエンドエフェクタ上で発生しているならば、適用が可能である。

双腕によるハンドルのステアリング操作、双脚によるクランクのペダリング操作ではともに、エンドエフェクタにおいて把持・接触が保たれているため、外的要因はその配下に存在

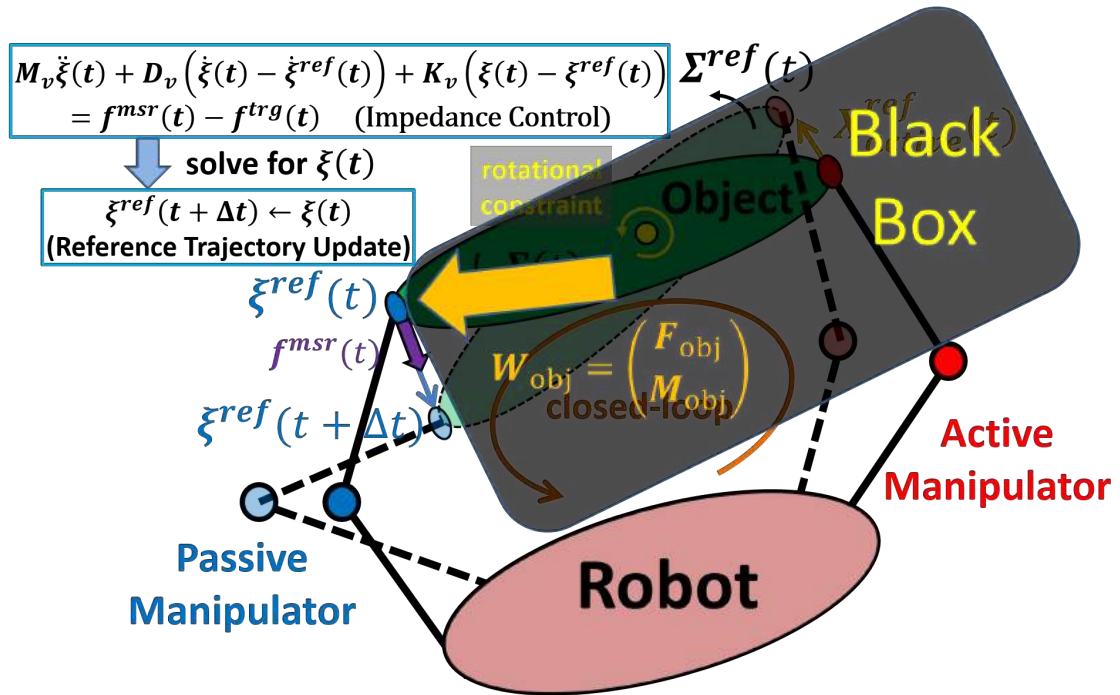


図 3.13: Control Law of Passive Manipulator with Black Box Object

する．さらに，双腕と双脚のマニピュレータ先端に6軸力センサを実装することで， \mathbf{W}_{obj} が計測される．これらの条件を満たすとき，マニピュレータに受動機能の制御が適用できるといえる．

位置指令型インピーダンス制御の適用

次にマニピュレータの運動（図 3.13 の左のマニピュレータ）に注目する．マニピュレータには今，外的要因 \mathbf{W}_{obj} がエンドエフェクタにおいて発生している．ここでマニピュレータに対して，インピーダンス制御 [131][132][49][50] を適用することを考える．外力および外モーメントをエンドエフェクタで受ける場合，インピーダンス制御が有効な手段である．本研究では，Nozawa ら [134] による位置指令型インピーダンス制御を採用する．以下にその制御則を示す．

$$\mathbf{M}_v (\ddot{\xi}(t) - \ddot{\xi}^{ref}(t)) + \mathbf{D}_v (\dot{\xi}(t) - \dot{\xi}^{ref}(t)) + \mathbf{K}_v (\xi(t) - \xi^{ref}(t)) = \mathbf{f}^{msr}(t) - \mathbf{f}^{trg}(t) \quad (3.13)$$

式 3.13 は、文献 [134] で示されている式を 3 次元の位置の成分について書き表した制御式である。 $\xi(t) \in \mathbb{R}^3$ は、時刻 t におけるマニピュレータのエンドエフェクタの位置を表し、目標位置は $\xi^{\text{ref}}(t) \in \mathbb{R}^3$ で表される。なお、式 3.13 は文献 [134] の加速度項に関して、目標値 $\xi^{\text{ref}}(t)$ を含めて表現している。 $f^{\text{msr}}(t) \in \mathbb{R}^3$ が、エンドエフェクタにおける外的要因 W_{obj} の計測接触力となる。また、 $M_v, D_v, K_v \in \mathbb{R}^{3 \times 3}$ はそれぞれ仮想慣性行列、仮想粘性行列、仮想剛性行列を表し、 $f^{\text{trg}}(t) \in \mathbb{R}^3$ は目標接触力である。

外的要因 W_{obj} を受ける環境下において、インピーダンス制御適用により、ある時刻 t ではエンドエフェクタ位置が、本来の目標位置 $\xi^{\text{ref}}(t)$ から適用後の位置 $\xi(t)$ へと遷移される。これは、エンドエフェクタから逆運動学により変換された関節空間においても、同様の遷移が実行されている。

オンライン軌道獲得則

$\xi^{\text{ref}}(t) \rightarrow \xi(t)$ の変化は、接触対象物体からの外的要因 W_{obj} を避けて緩和する働きがある。このことは、先行研究 [135] 等においてもインピーダンス制御の適応行動として実証されている。 $\xi(t)$ を次の時刻での目標位置とすることで、随時緩和行動が可能となり操作対象物体の運動に応じた追従動作が生成される。このときの軌道獲得則は、以下の更新式で表現される。

$$\xi^{\text{ref}}(t + \Delta t) := \begin{cases} \xi(t) & (\xi(t) \in \text{ReachabilityMap}) \\ \xi^{\text{ref}}(t) & (\xi(t) \notin \text{ReachabilityMap}) \end{cases} \quad (3.14)$$

$\xi(t)$ が Reachability Map の集合内に属するならば、次ステップでの目標位置 $\xi^{\text{ref}}(t + \Delta t)$ をインピーダンス制御式 3.13 適用後の現在位置 $\xi(t)$ としてセットする。これにより、受動機能のマニピュレータは接触対象の物体の動きに追従した軌道を逐次獲得できる。一方で、 $\xi(t)$ が Reachability Map の集合内に属さないならば、現在の目標位置 $\xi^{\text{ref}}(t)$ をそのまま保持する。

式 3.14 に基づいた逐次的なオンライン軌道獲得法は、制御入力を自動的に更新し決定していることに等しく、受動機能の制御則の根幹をなす。能動機能のマニピュレータの制御法では、ハンドルの長さなどの操作対象物の幾何的なモデルの情報が必要となるが、受動機能のマニピュレータの制御においてはこれらの情報を必要としないことが大きな特徴である。これは、従来のマスター・スレイブ型制御 [115] と異なる点である。また、受動機能の制御入力列は能動機能の制御則同様に、式 3.10 - 式 3.12 の条件を満たす必要がある。

オンライン探索指標を用いた定式化

位置指令型インピーダンス制御（式 3.13）とオンライン軌道獲得則（式 3.14）から、オンライン探索指標に基づいた一般化更新則を定式化する。位置指令型インピーダンス制御式 3.13 にラプラス変換を施し、 $\xi(t)$ のラプラス変換 $\xi(s)$ について解くと以下の式が得られる。

$$\xi(s) = \xi^{\text{ref}}(s) + (M_v s^2 + D_v s + K_v)^{-1} (f^{\text{msr}}(s) - f^{\text{trg}}(s)) \quad (3.15)$$

$\xi^{\text{ref}}(s), f^{\text{msr}}(s), f^{\text{trg}}(s)$ はそれぞれ $\xi^{\text{ref}}(t), f^{\text{msr}}(t), f^{\text{trg}}(t)$ のラプラス変換を表す。 $\xi(t)$ が Reachability Map 内に存在するとき、式 3.15 を時間領域に直した式を、式 3.14 の上式に代入することで $\xi^{\text{ref}}(t + \Delta t)$ と $\xi^{\text{ref}}(t)$ の関係が以下のように得られる。

$$\xi^{\text{ref}}(t + \Delta t) := \xi^{\text{ref}}(t) + \mathcal{L}^{-1} \left[(M_v s^2 + D_v s + K_v)^{-1} (f^{\text{msr}}(s) - f^{\text{trg}}(s)) \right] \quad (3.16)$$

ここで、軌道を $\rho(t; n) = \xi^{\text{ref}}(t)$ 、探索変数を $\mathbf{x}(t; n) = f^{\text{msr}}(t)$ としたときのオンライン制御指標における行列 $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ を式 2.6 と式 2.7 に従って求める。 $\rho(t; n) \in \mathbb{R}^3, \mathbf{x}(t; n) \in \mathbb{R}^3$ であることから、 $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ は $\mathbb{R}^{3 \times 3}$ の正方行列となる。

さて、式 2.7 の定義で示した通り、 $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ の各要素は軌道 $\rho_i(t; n)$ に対する探索変数 $x_j(t; n)$ の単調性から判断される。図 3.13 のようにマニピュレータのエンドエフェクタが操作対象と接触状態にあるとき、 $\rho_i(t; n) = \xi_i^{\text{ref}}(t)$ が増加するときの $x_j(t; n) = f_j^{\text{msr}}(t)$ の挙動は、軌道（目標位置）の成分と探索変数（力）の成分の座標系が揃っている場合 ($i = j$)、広義単調減少を呈し、式 2.7 の分類から $\text{sgn}[\rho_i(t; n), x_j(t; n)] = -1$ となる。このことは表 2.3 で示した行動例のように、目標となる軌道が増すほどエンドエフェクタと接触対象の間に働く内力負荷が徐々に増し、軌道とは逆の方向に力を受けることから定性的に観察できる。また、軌道の成分と探索変数の成分が異なる場合 ($i \neq j$)、互いに独立となり非依存に分類されることから $\text{sgn}[\rho_i(t; n), x_j(t; n)] = 0$ となる。したがって、これらの要素から構成される行列 $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ は以下で与えられる。 E は単位行列とする。

$$\text{sgn}[\rho(t; n), \mathbf{x}(t; n)] = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = -E \quad (3.17)$$

式 3.17 の $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ を用いて、式 3.16 で示した関係式を記述し直すと以下の軌道の更新則が定式化できる。

$$\xi^{\text{ref}}(t + \Delta t) := \xi^{\text{ref}}(t) - \mathcal{L}^{-1} \left[(M_v s^2 + D_v s + K_v)^{-1} \text{sgn}[\rho(t; n), \mathbf{x}(t; n)] (f^{\text{msr}}(s) - f^{\text{trg}}(s)) \right] \quad (3.18)$$

上式は式 2.14 で一般化したオンライン探索指標に基づく軌道修正の更新則の形式で表され、対応する各パラメータは以下の通りである。

表 3.3: 受動機能の制御法におけるオンライン探索指標および軌道修正のパラメータの対応

探索変数	探索目標
$f^{\text{msr}}(t)$	$f^{\text{trg}}(t)$
更新される軌道	一般化更新則の重み A
$\xi^{\text{ref}}(t)$	$(M_v s^2 + D_v s + K_v)^{-1}$

重み行列 A が正定値対称行列であることは、以下の点より自明である。

1. 仮想慣性行列 M_v 、仮想粘性行列 D_v 、仮想剛性行列 K_v はそれぞれ正定値対称行列である。
2. 正定値対称行列の逆行列もまた正定値対称行列である。

3.4.4 マニピュレータに対する探索指標の切替方策

双脚によるペダリング操作では、ユニラテラルな拘束性からクランクを引く力を作用できないため、オフライン探索指標に基づいた能動機能と、オンライン探索指標に基づいた受動機能を適宜切り替える必要がある。このような操作の切替行動は、バイラテラルな拘束をもつ双腕によるステアリング操作に対しても有効であり、マニピュレータ共通の探索指標の切替方策が提案されることが望まれる。

具体的に三輪車の搭乗型機器の例では、双腕と双脚の操作行動における機能の切替は以下のような場合に実行されるとよい。

- 双脚によるペダリング操作での、クランク角の切替角到達時 ($\theta_c = \frac{\pi}{2} + \delta_c, \frac{3\pi}{2} + \delta_c, \dots$) における踏み込み期の脚の切り替え
- 双腕によるステアリング操作での、任意のタイミングにおける切り替え

ハンドル・クランクの運動計測

能動と受動の各探索指標の機能切替は、操作対象物の運動状況に応じて行う。一般に三輪車のハンドルやクランクといった 1 自由度の回転拘束をもつ剛体では、角度・角速度などが

運動計測のためのパラメータとなる。これらのパラメータはヒューマノイドが搭乗する三輪車のハンドル・クランクに角度取得のためのエンコーダや、角速度計測のためのジャイロセンサを取り付けることで計測可能であるが、本研究においては三輪車側に計測センサを設けず、ロボット搭載のセンサで計測することとする。ヒューマノイドの内界センサ・外界センサを利用したこれらのパラメータ計測手法として、

(a) ロボットの運動学モデルと関節エンコーダ情報に基づいた手先足先位置による操作対象物（ハンドル・クランク）の角度の把握 [136]

(b) オプティカルフロー等のビジョンを利用したハンドルの角速度の実時間推定 [137] などが挙げられる。

(b) の手法は搭乗するヒューマノイドの可視領域に存在するハンドルのみにも適用可能であるため、ハンドル・クランク共通の運動計測方法として、本論文では手法 (a) を採用してハンドル・クランクの角度情報を運動状況とみなす。手法 (b) をハンドルの運動状況の把握のみに適用した場合、得られる角速度情報を利用したハンドル操作の調整に基づく切替則が定義できる。例として、ハンドル操舵における所望の回転角速度に応じたヒューマノイド双腕の能動機能・受動機能の切替制御法が考察される。

能動機能と受動機能における探索指標の切替則

前項で採用した手法 (a) による運動計測に基づき、操作対象物であるハンドルやクランクの角度 θ_h, θ_c について、双腕と双脚のマニピュレータの機能のオフライン探索指標とオンライン探索指標の切替点（切替角度） $\theta_h^{SW}, \theta_c^{SW}$ を以下の式により定める。

$$\theta_*^{SW} = \theta_*^{init} + k\pi, \quad * = h \text{ or } c, \quad k \in \mathbb{Z}(\text{integer}) \quad (3.19)$$

$\theta_h^{init}, \theta_c^{init}$ はハンドルやクランクの初期切替角を表し、例えばハンドルの初期切替角の場合は式 3.2 で示した θ_h の取りうる範囲において、2つのマニピュレータの能動・受動機能を等しく分担できる中央値として $\theta_h^{init} = 0$ 、クランクの初期切替角の場合は式 3.6 より $\theta_c^{init} = \frac{\pi}{2} + \delta_c$ と設定できる。また、 k は \mathbb{Z} （整数集合）の範囲の離散値を取りうる。手法 (a) により得られた三輪車のハンドル角・クランク角 θ_h, θ_c が、式 3.19 から得られる切替点 $\theta_h^{SW}, \theta_c^{SW}$ を通過するとき、2つのマニピュレータの探索指標（能動機能におけるオフライン探索指標と受動機能におけるオンライン探索指標）の機能切替が実行される。以上の切替則にしたがった双腕と双脚に対する探索指標の切替判断システムを図 3.14 に示す。

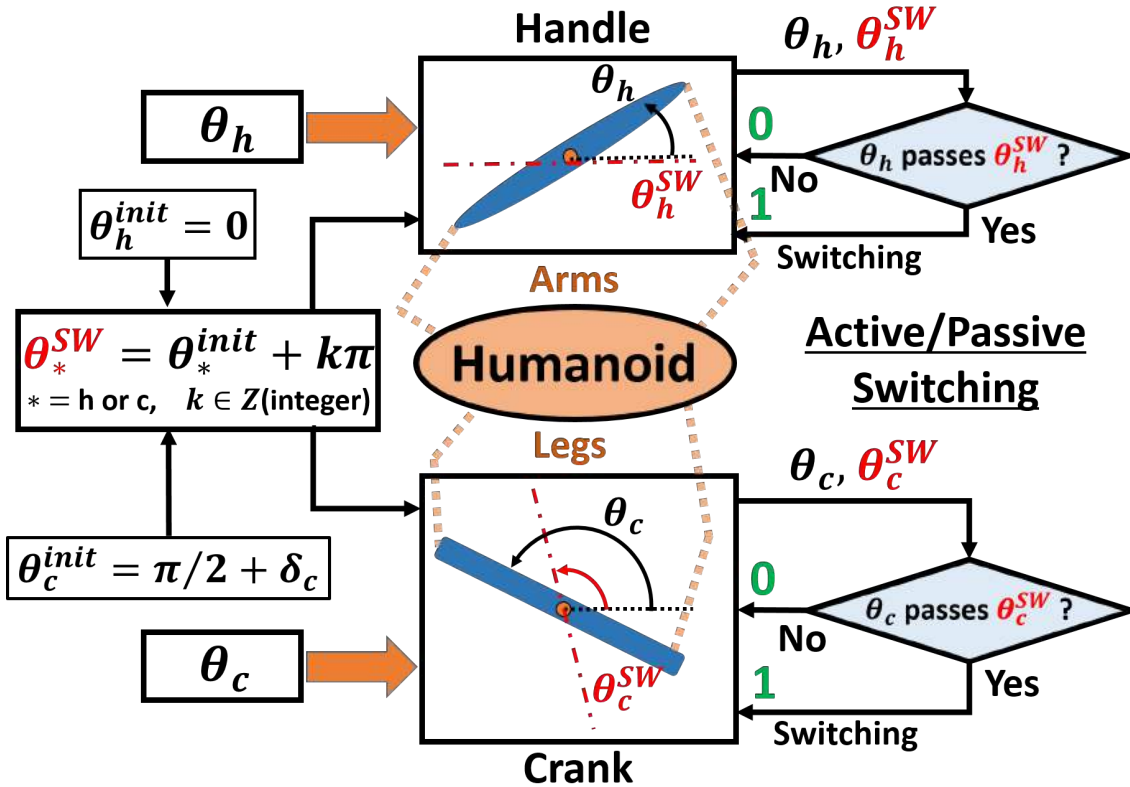


図 3.14: Active/Passive Switching Diagnostic System for Arms and Legs of Humanoid Robot Based on Angles of Handle and Crank [126]

3.5 ヒューマノイドの双腕・双脚による搭乗型機器の運転操作

探索指標の考えをもとに前節で一般化した能動機能と受動機能の制御法を、ヒューマノイドによる三輪車操作行動へと適用し、制御則の妥当性と有用性を実証する。本実機検証では、運転操作行動の実験環境として室内環境および屋外環境の2つの環境で実施し、等身大ヒューマノイドロボット HRP2-JSK [110] の他、筋骨格腱駆動ヒューマノイド“腱志郎”[138][139] の2台の異なる身体構成を有するヒューマノイドロボットをそれぞれ用いた。

3.5.1 実験準備

図 3.15 のように等身大ヒューマノイドロボット HRP2-JSK [110] を三輪車に搭乗させた状態を構成する。三輪車のタイヤと実験環境の路面は十分な静止摩擦が保たれており、スリップやドリフトなどの滑り運動がないものとする。

3.4 節で提案した探索指標に基づく能動機能と受動機能の制御法を三輪車操作に適用させた場合の制御システム構成を図 3.16 に示す。三輪車のハンドル角指令 θ_h 、クランク角指令 θ_c を制御入力として与え、本実験ではそれぞれの制御入力のサンプリング間隔を $\Delta\theta_h = 5[\text{deg}]$, $\Delta\theta_c = 10[\text{deg}]$ とする。

図 3.16 に示した操作行動の制御システムにおける一連の動作生成アルゴリズムは次の通りである。

1. 三輪車の運動学モデルに対し、目標ハンドル角 θ_h ・目標クランク角 θ_c を入力
2. 目標角を実現する逆運動学計算による能動機能の動作生成 (Δt で遷移)
3. 実ハンドル・クランクに追従するような受動機能の動作生成 (Δt で遷移)
4. 目標ハンドル角 θ_h ・目標クランク角 θ_c の更新

$$(\theta_h \leftarrow \theta_h \pm \Delta\theta_h, \theta_c \leftarrow \theta_c \pm \Delta\theta_c)$$

適用する各実験パラメータは、表 3.4 の通りである。

表 3.4: Experimental Parameters (Impedance Controller Parameters, Target Force Parameters, and Max Handle Angle) for Active Arm, Passive Arm, Active Leg, and Passive Leg

	Active Arm	Passive Arm	Active Leg	Passive Leg
M_p [$\text{N} \cdot \text{s}^2/\text{m}$]	10.0	10.0	10.0	10.0
D_p [$\text{N} \cdot \text{s}/\text{m}$]	600.0	500.0	600.0	500.0
K_p [N/m]	3000.0	300.0	3000.0	300.0
f^{trg} [N]	$(0.0 \ 0.0 \ 0.0)^T$	$(0.0 \ 0.0 \ 0.0)^T$	$(0.0 \ 0.0 \ 150.0)^T$	$(0.0 \ 0.0 \ 150.0)^T$
Θ_H [deg]	25.0	25.0	-	-

ここでは、動的かつ接触対象拘束な閉ループ系における各マニピュレータへの過負荷防止を図るため、能動機能のマニピュレータには高剛性のインピーダンス制御パラメータ ($M_p =$



図 3.15: Life-Sized Humanoid Robot HRP2-JSK [110] Riding Tricycle

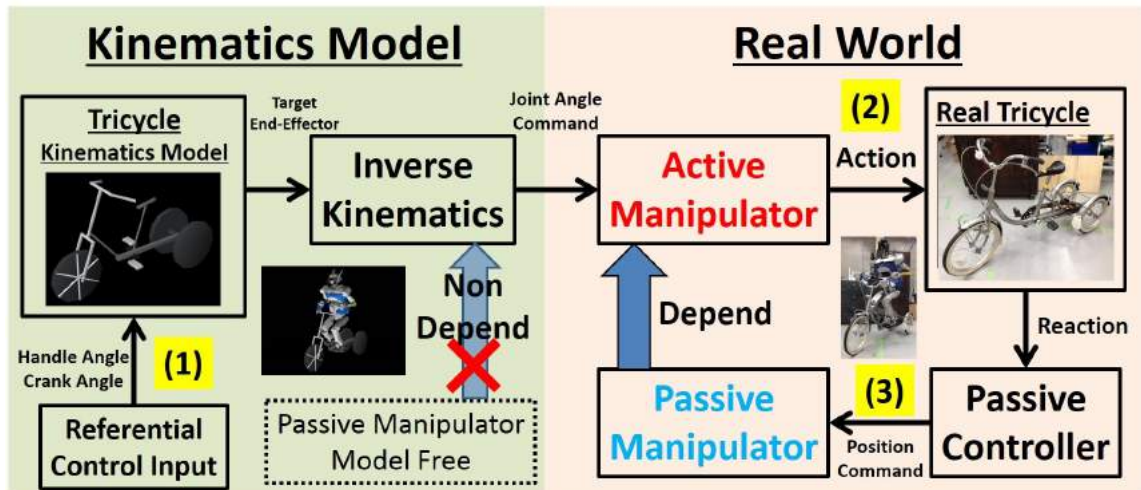


図 3.16: Tricycle Control System on Experiment by Humanoid Robot [52]
 Command from kinematics model to real world.

10, $D_p = 600$, $K_p = 3000$), 受動機能のマニピュレータには低剛性のインピーダンス制御パラメータ ($M_p = 10$, $D_p = 500$, $K_p = 300$) を導入した. M_p, D_p, K_p はそれぞれ, 位置に対する慣性パラメータ, 粘性パラメータ, 剛性パラメータである.

さらに, 足部とペダルのようにピンディングが設けられていないユニラテラルな拘束の場合, マニピュレータと接触対象物との間に一定の力を作用させる必要があり, 接触力が低下すると摩擦力の低下ならびに接触部位の離脱が問題となる. そのため, 本研究ではユニラテラル性を呈する双脚のペダリング操作において, 両足裏-両クランクペダル間に一定の目標接触力を導入し, $f^{trg}(t) = (0 \ 0 \ 150)^T$ [N] とした. これにより, 漕ぎ操作において大きく寄与する鉛直 (z 成分) 接触力の低下防止を目標とする.

3.5.2 操作行動の実証実験

ヒューマノイドの双腕と双脚の組に, オフライン探索指標に基づく能動機能とオンライン探索指標に基づく受動機能の軌道制御を導入し, ステアリング操作とペダリング操作の両タスクを実現する. また, これらの両タスクを統合した双腕と双脚によるカーブ走行・蛇行走行についても達成させる.

双腕によるステアリング操作

図 3.17 に双腕によるステアリング操作の行動実験結果を示す。ハンドル回しの操舵範囲は式 3.2, 式 3.4 に基づき, $-25[\text{deg}] \leq \theta_h \leq 25[\text{deg}]$ とした。 $-25[\text{deg}] \leq \theta_h \leq 0[\text{deg}]$ の範囲は, 右腕をオンライン探索指標を適用した受動機能, 左腕をオフライン探索指標を適用した能動機能とした。一方, $0[\text{deg}] \leq \theta_h \leq 25[\text{deg}]$ の範囲は, 右腕をオフライン探索指標を適用した能動機能, 左腕をオンライン探索指標を適用した受動機能とする。両探索指標に対するハンドルの切替角は, 式 3.19 に基づき $\theta_h^{\text{init}} = 0[\text{deg}]$ としている。

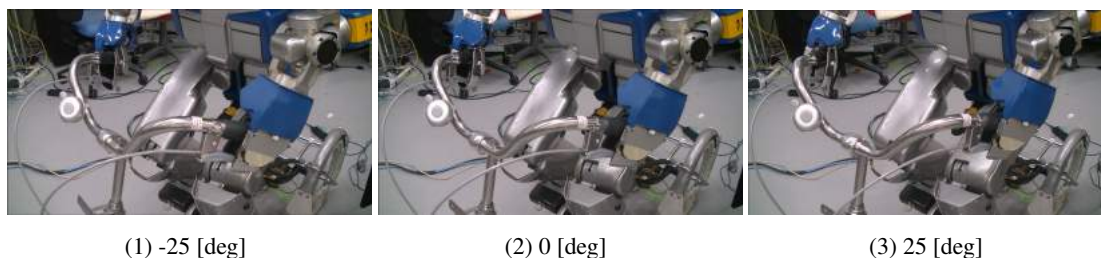


図 3.17: $-25[\text{deg}] \leq \theta_h \leq 0[\text{deg}]$: Left Arm is Active, and Right Arm is Passive
 $0[\text{deg}] \leq \theta_h \leq 25[\text{deg}]$: Right Arm is Active, and Left Arm is Passive

双脚によるペダリング操作

クランクのペダリング操作ではユニラテラルな拘束性から, クランクの上死点から下死点まで踏み下ろす脚にオフライン探索指標を用いた能動機能, 下死点から上死点まで引く脚にオンライン探索指標を用いた受動機能を適用する。すなわち, 1 回転漕ぐような条件式 3.5 を満たすクランク角の範囲において,

- $0 \leq \theta_c < \frac{\pi}{2} + \delta_c$, $\frac{3\pi}{2} + \delta_c \leq \theta_c < 2\pi$ のとき,
 左脚 (踏み込み脚) をオフライン探索指標による能動機能, 右脚 (引き脚) をオンライン探索指標による受動機能
- $\frac{\pi}{2} + \delta_c \leq \theta_c < \frac{3\pi}{2} + \delta_c$ のとき,
 右脚 (踏み込み脚) をオフライン探索指標による能動機能, 左脚 (引き脚) をオンライン探索指標による受動機能

とする。クランクの各探索指標の切替角は図 3.7 [125] にしたがって、 $\delta_c = 20[\text{deg}]$ とした。したがって、式 3.19 の初期切替角は $\theta_c^{\text{init}} = \frac{\pi}{2} + \delta_c (= \frac{\pi}{2} + 20 \cdot \frac{\pi}{180})[\text{rad}]$ である。以上をクランクの操作の死点として、能動機能と受動機能の探索指標の切替をおこなった (図 3.18)。

クランク回し 1 周 ($0 \leq \theta_c < 2\pi$) においてかかる時間 $T_{2\pi}$ は、クランク角 $\Delta\theta_c (= 10 \cdot \frac{\pi}{180}[\text{rad}])$ の遷移に対して $2\Delta t$ (= 能動機能の脚の動作遷移時間 Δt + 受動機能の脚の動作遷移時間 Δt) がかかることから、

$$T_{2\pi} = \frac{2\pi \cdot 2\Delta t}{\Delta\theta_c} = \frac{2\pi \cdot 2 \cdot 0.8}{10 \cdot \frac{\pi}{180}} = 57.6[\text{s}] \quad (3.20)$$

と見積ることができる。本実験における各機能の動作遷移時間 Δt はヒューマノイドの関節モータへの負荷の影響を考慮して $\Delta t = 800[\text{ms}] (= 0.8[\text{s}])$ と設定している。

双腕・双脚による統合操作

双腕によるステアリング操作と双脚によるペダリング操作を組み合わせたカーブ走行・蛇行走行を示す。図 3.19, 図 3.20, 図 3.21 はそれぞれ、

- 操舵角を下限 $\theta_h = -\Theta_H = -25[\text{deg}]$ としたカーブ走行
- 操舵角を上限 $\theta_h = \Theta_H = 25[\text{deg}]$ としたカーブ走行
- $\theta_h = -\Theta_H = -25[\text{deg}]$ (if $0[\text{deg}] \leq \theta_c \leq 180[\text{deg}]$),
 $\theta_h = \Theta_H = 25[\text{deg}]$ (if $180[\text{deg}] \leq \theta_c \leq 360[\text{deg}]$) とした蛇行走行

の実験結果を表している。

以上の実証実験の結果より、提案した探索指標の考えに基づく制御法をヒューマノイド実機に適用することで、搭乗型機器の操作行動が適切に実現されている。

3.5.3 三輪車操作実験における能動・受動の機能考察

能動機能と受動機能の比較及び考察

前の実験結果より、能動機能と受動機能の特性について考察を行う。双脚によるペダリング操作において、必要条件となる両足のクランクペダルに対する目標接触力への追従について、能動機能の場合と受動機能の場合を比較し考察する。ここでは、双脚を常に能動機能として作用させた場合と、双脚に能動機能と受動機能の両特性を備えた場合の漕ぎ操作を比較する。



(1) Active: Left Leg
Passive: Right Leg



(2) Active: Left Leg
Passive: Right Leg



(3) Active: Left Leg → Right Leg
Passive: Right Leg → Left Leg



(4) Active: Right Leg
Passive: Left Leg



(5) Active: Right Leg
Passive: Left Leg



(6) Active: Right Leg → Left Leg
Passive: Left Leg → Right Leg



(7) Active: Left Leg
Passive: Right Leg



(8) Active: Left Leg
Passive: Right Leg

図 3.18: Pedaling Manipulation on 1-Cycle by Active Leg and Passive Leg



(1) Active: Left Leg
Passive: Right Leg



(2) Active: Left Arm
Passive: Right Arm



(3) Active: Left Leg
Passive: Right Leg



(4) Active: Left Leg → Right Leg
Passive: Right Leg → Left Leg



(5) Active: Right Leg
Passive: Left Leg



(6) Active: Right Leg
Passive: Left Leg



(7) Active: Right Leg → Left Leg
Passive: Left Leg → Right Leg



(8) Active: Left Leg
Passive: Right Leg

図 3.19: Steering and Pedaling Manipulation (-25[deg] curve) by Active/Passive Arms and Legs



(1) Active: Left Leg
Passive: Right Leg



(2) Active: Right Arm
Passive: Left Arm



(3) Active: Left Leg
Passive: Right Leg



(4) Active: Left Leg → Right Leg
Passive: Right Leg → Left Leg



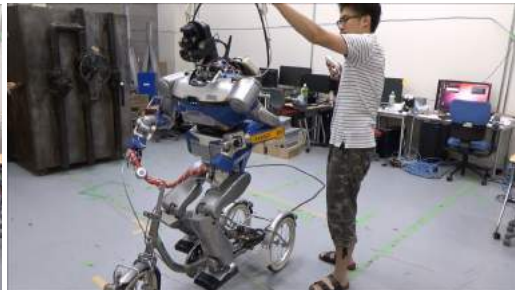
(5) Active: Right Leg
Passive: Left Leg



(6) Active: Right Leg
Passive: Left Leg



(7) Active: Right Leg → Left Leg
Passive: Left Leg → Right Leg



(8) Active: Left Leg
Passive: Right Leg

☒ 3.20: Steering and Pedaling Manipulation (25[deg] curve) by Active/Passive Arms and Legs

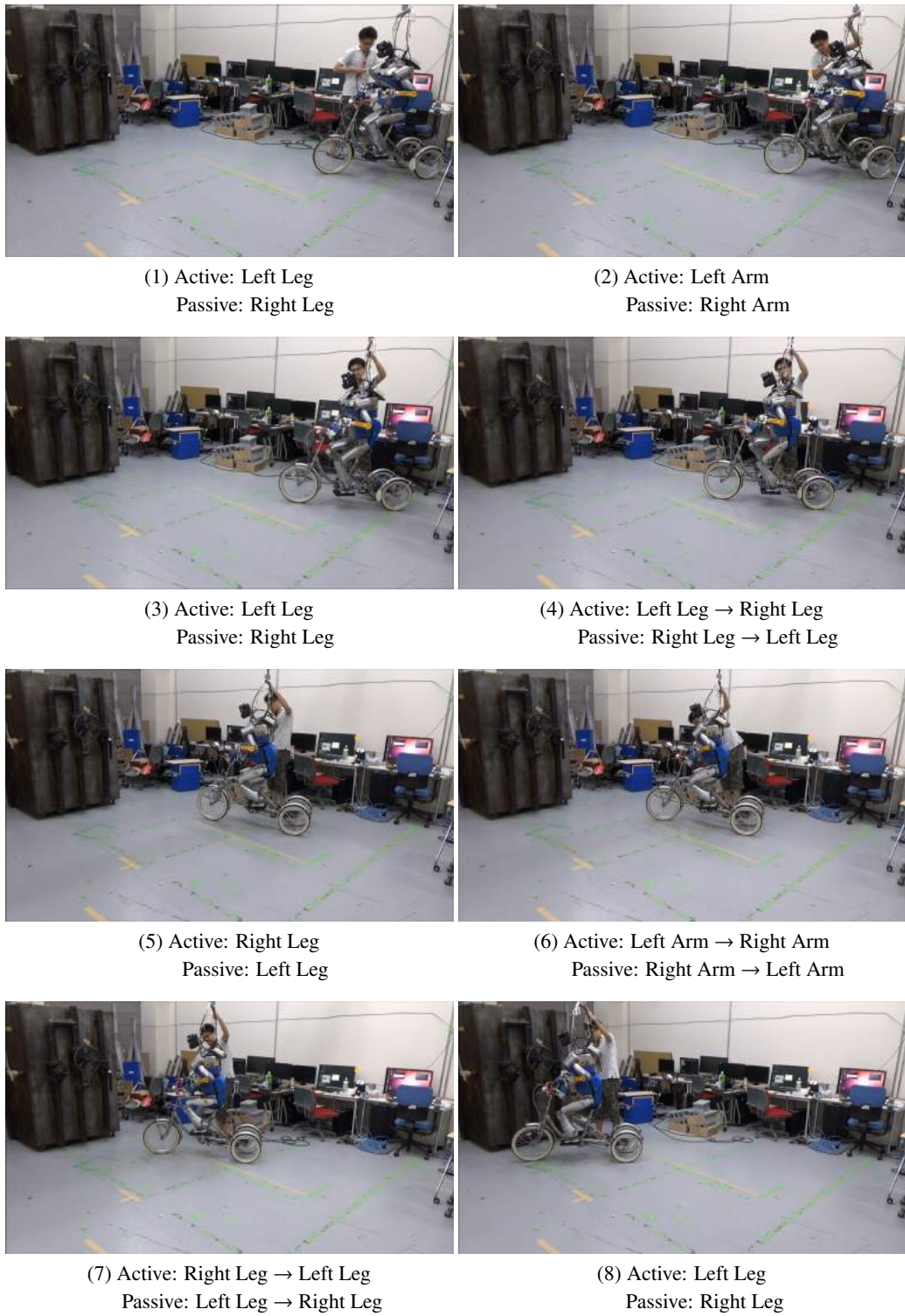


図 3.21: Steering and Pedaling Manipulation (Switching Handle from -25[deg] to 25[deg]) by Active/Passive Arms and Legs

このときの両足裏(左足・右足)にかかる鉛直成分の計測接触力 f_z の推移は図 3.22 となる. 図 3.22 -(I) が双脚とも能動機能である場合であり, 図 3.22 -(II) が提案した能動機能と受動機能の制御則を双脚に適用した場合の結果となる. 縦の黒線が切替角の到達時刻である.

比較の結果, ともに目標接触力の鉛直成分は $f^{trg}(t)$ で示した 150[N] を採用しているが, 双脚とも能動機能の場合は両足の目標接触力に対する追従が難しく, 160[N] 付近を中心に高低変動が観察された. 片方の脚マニピュレータに受動機能を備えることで, 図 3.22 -(II) のように目標接触力追従が可能となり, 受動機能側の足とペダルの接触力が十分に保たれていることが分かる.

ここで示した能動機能と受動機能の特性の違いは, 参照とする軌道が, オフラインで計画されたエンドエフェクタの位置軌道 $\mathbf{X}_{active}^{ref}(t)$ であるか, 式 3.18 からオンラインに獲得された追従軌道 $\xi^{ref}(t)$ であるか, の違いによるものと推察できる. 双脚はともにインピーダンス制御(式 3.13) を適用しているが, 能動機能の場合は, 計画した軌道 $\mathbf{X}_{active}^{ref}(t)$ とインピーダンス制御式 3.13 により得られる位置 $\xi(t)$ の残差が生じることで力の追従誤差が発生している. 一方で受動機能の場合は, インピーダンス制御式 3.13 で得られた位置 $\xi(t)$ をそのまま次ステップの追従軌道 $\xi^{ref}(t + \Delta t)$ として獲得することで, 目標接触力 $f^{trg}(t)$ への追従を反映した軌道生成が実現できていることが考察される.

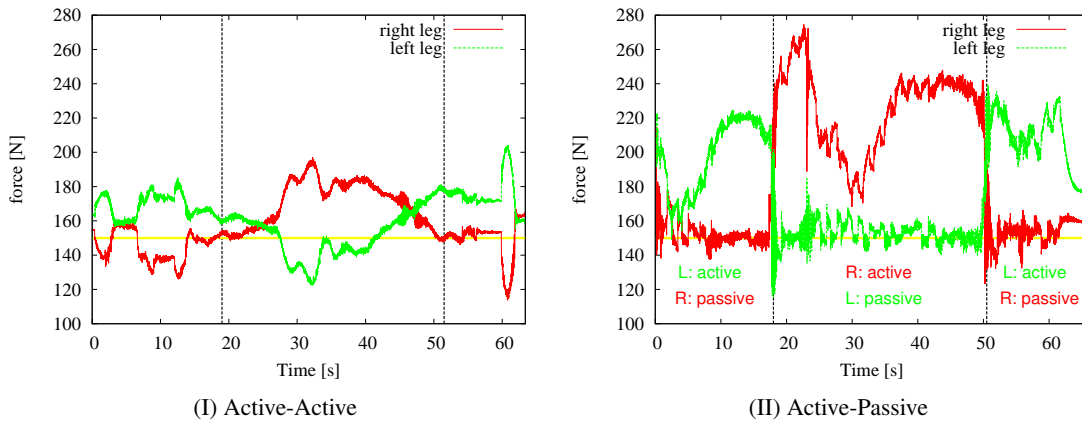


図 3.22: Graphs Comparing Only-Active and Active-Passive about Vertical Force Change on Feet of Humanoid during Pedaling [53]

能動機能側の接触力の増大に関する議論

先ほど議論した受動機能の目標接触力追従の特性と異なり, 能動機能を備えた脚については図 3.22 -(I),(II) とともに 150[N] を上回っている. 能動機能のマニピュレータの場合, 目標力に加えてエンドエフェクタの参照軌道への追従動作が実行されるため, 大きな漕力が提供されることでクランクの回転運動へと伝達される. 特に図 3.22 -(II) で示した能動機能側の脚の鉛直接触力を観察すると, 上死点及び下死点付近 (縦の黒線付近) においては鉛直接触力の増大が見られ, クランクが水平となる地点 (図 3.22 -(II) の 3[s], 30[s], 63[s] 付近) では下回って 150[N] に近づくような波形が見られる.

この現象は, 目標力ベクトルを $f^{trg}(t) = (0 \ 0 \ 150)^T$ [N] としして鉛直方向に一定に設けている一方で, エンドエフェクタの追従軌道のベクトルがクランク角に依存した接線ベクトルであることから推察できる. 図 3.22 -(II) の 3[s], 30[s], 63[s] 付近のクランクが水平となる地点では, 目標力ベクトルの方向とクランク接線方向がともに鉛直方向で一致し, 目標力である 150[N] に近づく. 反対に, 上死点・下死点に近づくとき目標力の方向 (鉛直方向) と軌道追従すべき方向 (接線方向) が大きくずれてくる. このとき, 参照軌道方向に対して本来必要とされる 150[N] を満たすように鉛直接触力が目標値 150[N] を超えて増加する. 増加率は鉛直接触力と接線方向の関係から $1/|\cos \theta_c|$ のオーダーとなる. このオーダーが波形を決定する要因であることが考察できる.

能動機能側の接触力の増大を抑えたい場合, 鉛直 (z 成分) 接触力の低下防止に加えて, エンドエフェクタが追従する参照軌道の方向性についても目標接触力 $f^{trg}(t)$ において考慮することが課題となる.

3.5.4 実践的な運転操作行動の実現

本項では, 実践を意識したヒューマノイドによる三輪車操作行動の実現を示す. 人間が指令する操作のためのインターフェースシステムを構築し, 操作インターフェースシステムに基づいた運転操作実験を室内環境と屋外環境の2種類の環境で実機検証した.

操作インターフェースシステムの構築

ヒューマノイドが搭乗型機器の操作行動を実現するためには, 人間が明示的に指示を与える層 (指示層) と, 人間の非指示でヒューマノイド自身が行動を決定し生成する層 (自律層)

に分別し，行動の層・立ち位置を明確化することが重要である．ここでは前者の指示層に注目した搭乗型機器の運転操作を目的とした実践的な操縦のインターフェース構築を行う．

ヒューマノイドを行動させるために必要な指示方法として，デバイス入力による行動指示をおこなう．入力のためのデバイスは多種多様に存在するが，本研究では人間が意図して指示入力を与えるデバイスとして，ジョイスティックコントローラを用いたハンドル角・クラック角の目標角度指令を与える．ジョイスティックコントローラからの入力指令と，操縦用の計算機およびロボット体内の計算機への通信処理はROS(Robot Operating System) [140][141]が提供する通信ライブラリを使用する．詳細な低レイヤーにおけるソフトウェアシステムの構成についてはA.2.1項で述べるものとして，本章で提案したオフライン探索指標とオンライン探索指標の協調によるマニピュレータの能動機能と受動機能をベースとした操作インターフェースシステムを図3.23に示す．

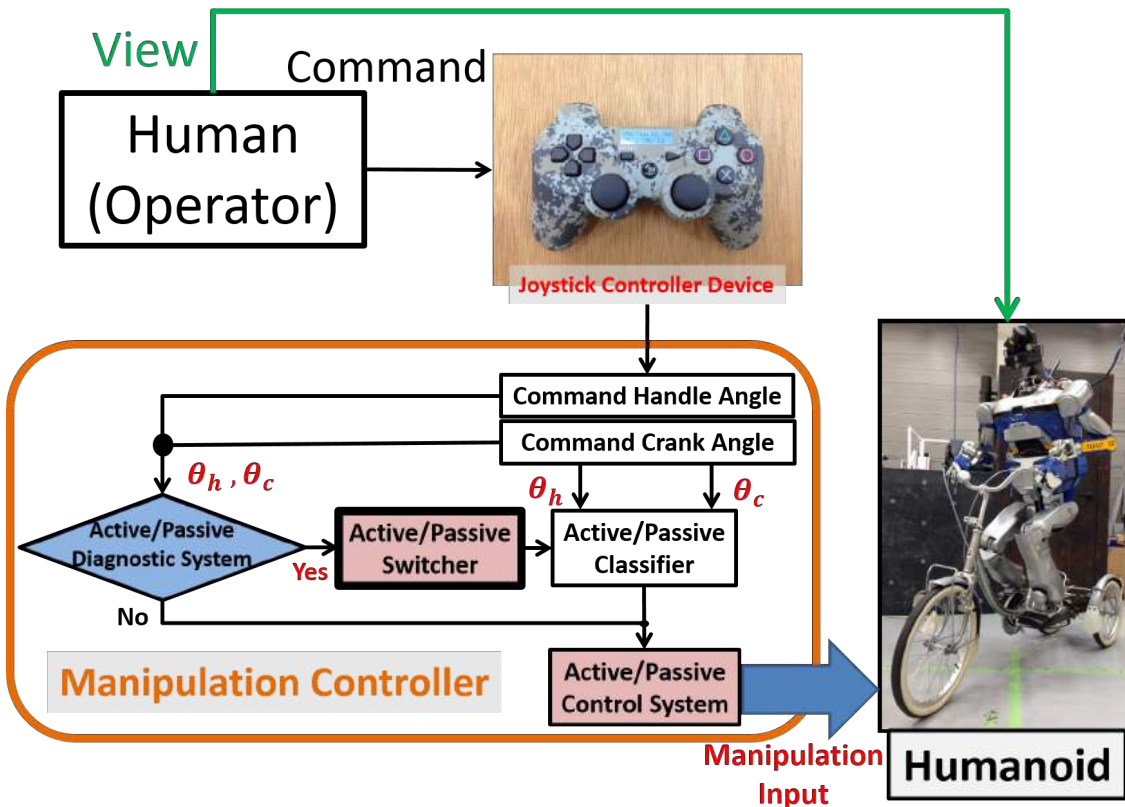


図 3.23: Operating Interface System by Joystick Controller Device

図 3.23 の操作システムは人間がロボット実機や周囲の外界環境を目視で観察した上で、その情報に基づき人間が所望の搭乗型機器への目標量（三輪車の場合は、目標ハンドル角と目標クランク角）を決定し、ジョイスティックコントローラに入力を与えることで運転操作行動を実現するシステム構成となっている。このシステムにおいては、人間からの指示は搭乗型機器への目標量のみで、ヒューマノイドの手先・足先の位置指令や目標関節角度への指令は与えずに、提案したマニピュレータの軌道修正法に基づきヒューマノイド自身が搭乗型機器への作用を決定し与えている。

実践的なヒューマノイドによる運転操作行動実験

前項で構成した運転操作のためのインターフェースシステムを利用し、室内環境および屋外環境でのヒューマノイド実機による実践的な運転操作行動の実現を行った。以下にそれぞれの実証実験について述べる。

室内環境でのコントローラデバイス指令に基づく運転操作行動実験 人間からの指示層であるジョイスティックコントローラのデバイス入力から三輪車のハンドル角 θ_h とクランク角 θ_c の指令角度を与えた運転操作行動を室内の平坦な床面上で検証した（図 3.24）。

本検証実験から以下のシステム・体系に関する有用性が示された。

- 構築した操作インターフェースシステム（図 3.23）の実機運転に対する有用性
- オフライン探索指標とオンライン探索指標の2つの探索指標の概念を交互・切替利用したマニピュレータの能動機能と受動機能の制御系の有用性

屋外環境でのコントローラデバイス指令に基づく運転操作行動実験 実験環境を均一な床面を有する室内環境から、不規則な凹凸路面を呈する屋外環境へと発展させる。室内環境実験と同様の制御システム、操作インターフェースシステムに基づく屋外環境下における実践的な運転操作行動を図 3.25、図 3.26 に示す。それぞれ、図 3.25 が遠方真横からみた走行実験の様子であり、図 3.26 が前近傍からみた運転操作のキャプチャである。約 12.8[m] にわたる凹凸路面の実環境を、ペダリング操作による進行ならびにステアリング操作による操舵を組み合わせながら、走破可能であることを実証した。

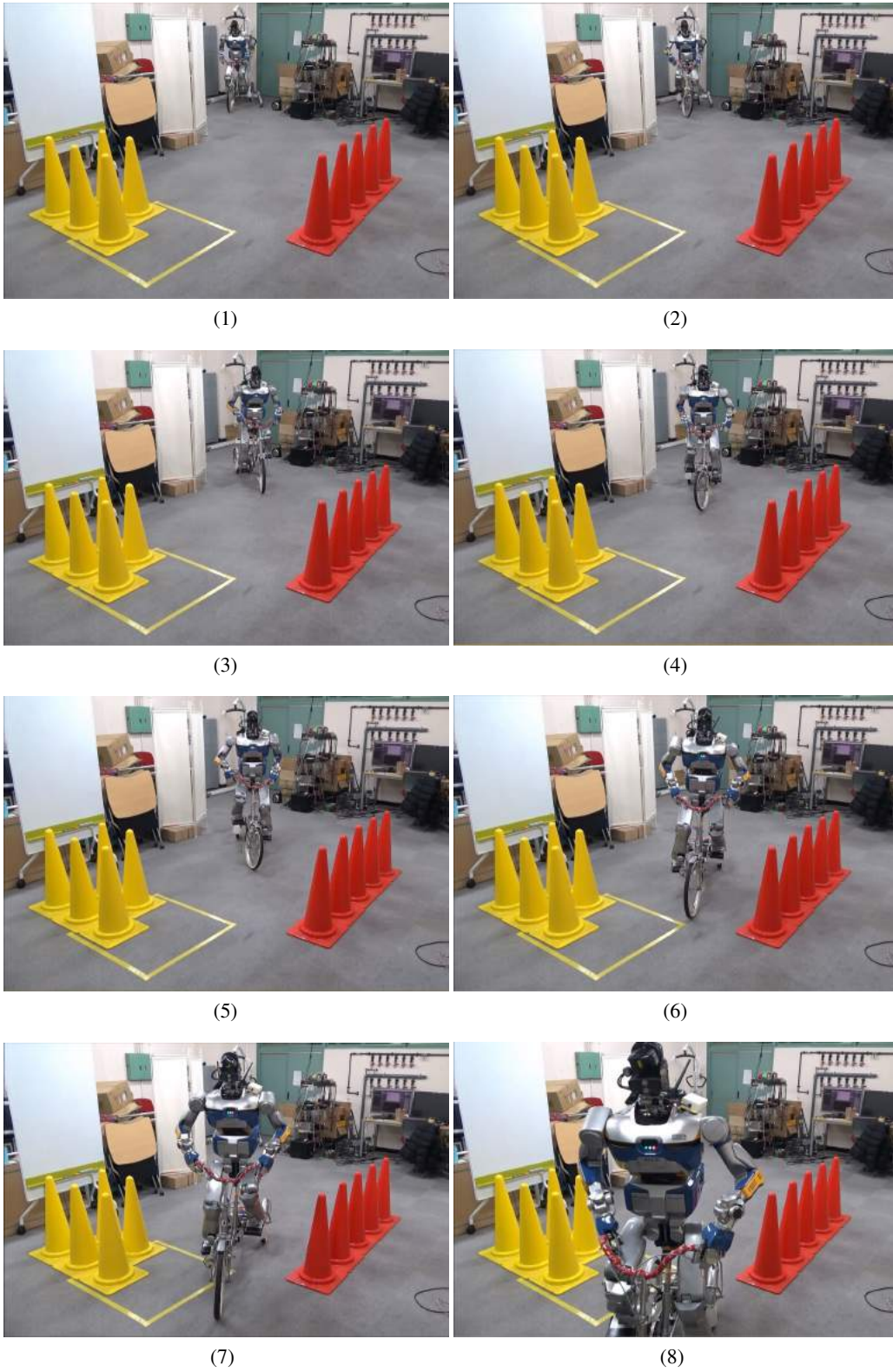


図 3.24: Slalom Steering and Pedaling by Joystick Controller Device

本実験の運転結果として、等身大ヒューマノイドロボット HRP2-JSK [110] による三輪車操作において、12.8[m] の屋外路面環境を 2 分 44 秒で走破した。このとき観察された走行の遅延性について考察する。操作行動のための制御システム（図 3.16）における 4 つのアルゴリズムで示したように、1 サンプリングの角度指令 θ_h, θ_c を実現するまでに図 3.16 内の (2) と (3) のフェーズで合計 $2\Delta t$ の時間がかかっている。したがって、(2) と (3) の間に反応の時間差が存在しており、発進（(1) と (2) のフェーズ）と停止（(3) と (4) のフェーズ）を繰り返した走行動作となっている。このことが運転操作行動に対する走行速度の遅延性を発生させている。また、上記の点に加えて漕力の提供のタイムラグも一因となり、足部-ペダル間の内力の蓄積による重心の Roll まわりの回転モーメントの外乱発生による転倒も懸念事項である。

走行の遅延性を改善した屋外環境下における運転操作行動実験 先に述べた屋外環境での運転操作行動実験から得られた走行の遅延性の考察・原因の推察を基に、改善方策を与えた屋外環境下での運転操作行動の再実験を実施した。前回の実験から改善したポイントを以下に整理する。

- ロボット実機のソフトウェアシステム¹⁾におけるヒューマノイドの目標関節角度までの遷移の最小時間のリミット引数“min-time”による制限を解除。
- 目標関節角度の指令に対して実機の動作遷移が終了するまでの待機時間を前回の待機時間の 80% まで短縮。
- オンライン探索指標に基づく受動機能の脚の目標接触力 $f^{trg}(t)$ を $(0\ 0\ 150)^T$ [N] から $(0\ 0\ 50)^T$ [N] に下げ、オフライン探索指標による能動機能とオンライン探索指標による受動機能の間の目標接触力の差を大きくすることで、ペダリング操作におけるクランクの回転トルクを増大。
- クランクの能動機能と受動機能の切替角 δ_c のパラメータ調整（⇒ 走行スピードを最大化させる切替角の最適値の探索が望まれる）。
- ハンドルのステアリング操作行動生成とクランクのペダリング操作行動生成の間の動作遷移の待機を無くし、両操作行動を同時に生成できるように改善（ハンドルを切っ

¹⁾https://github.com/jsk-ros-pkg/jsk_pr2eus

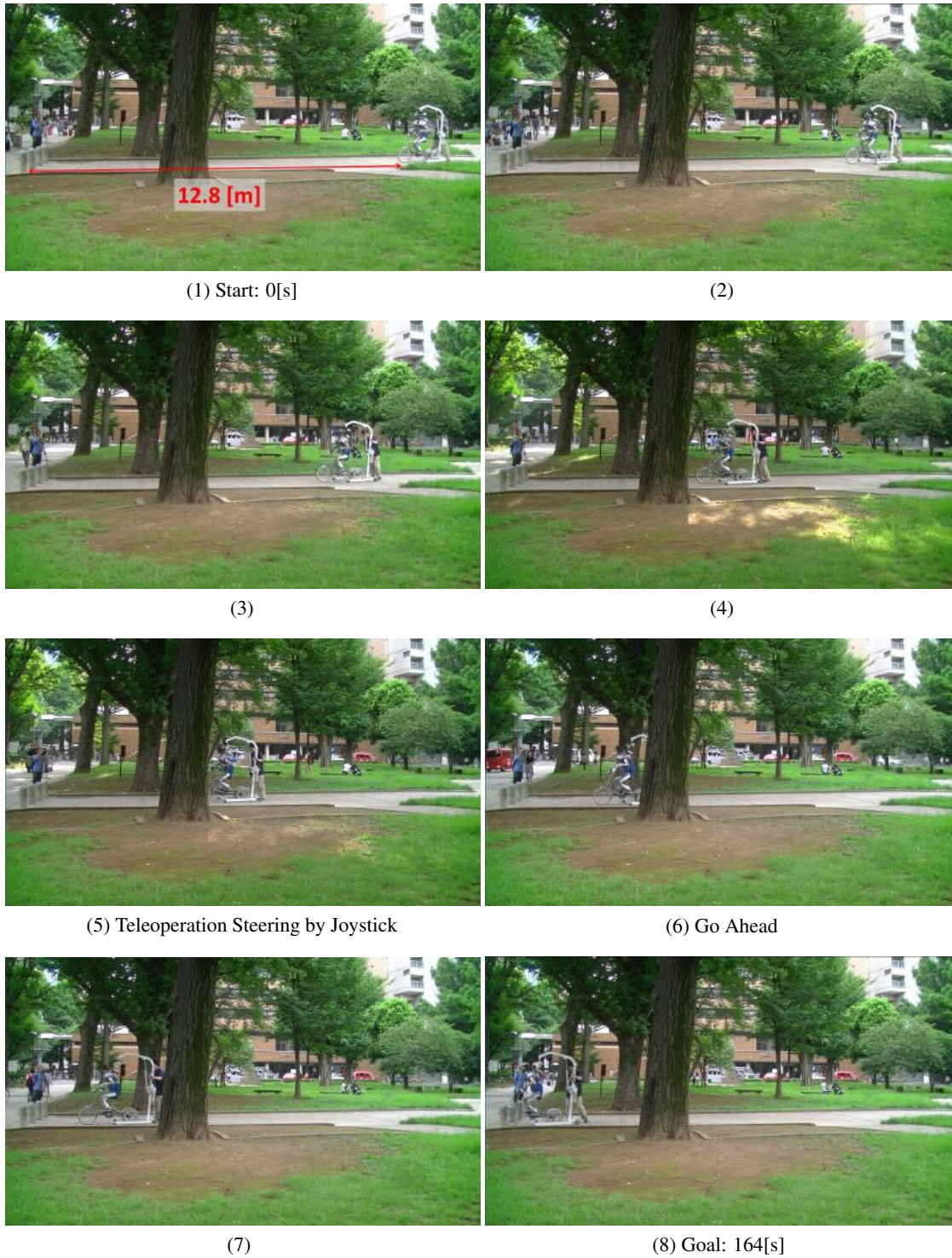


図 3.25: Steering and Pedaling by Joystick Controller Device in the Open Air (Side View)

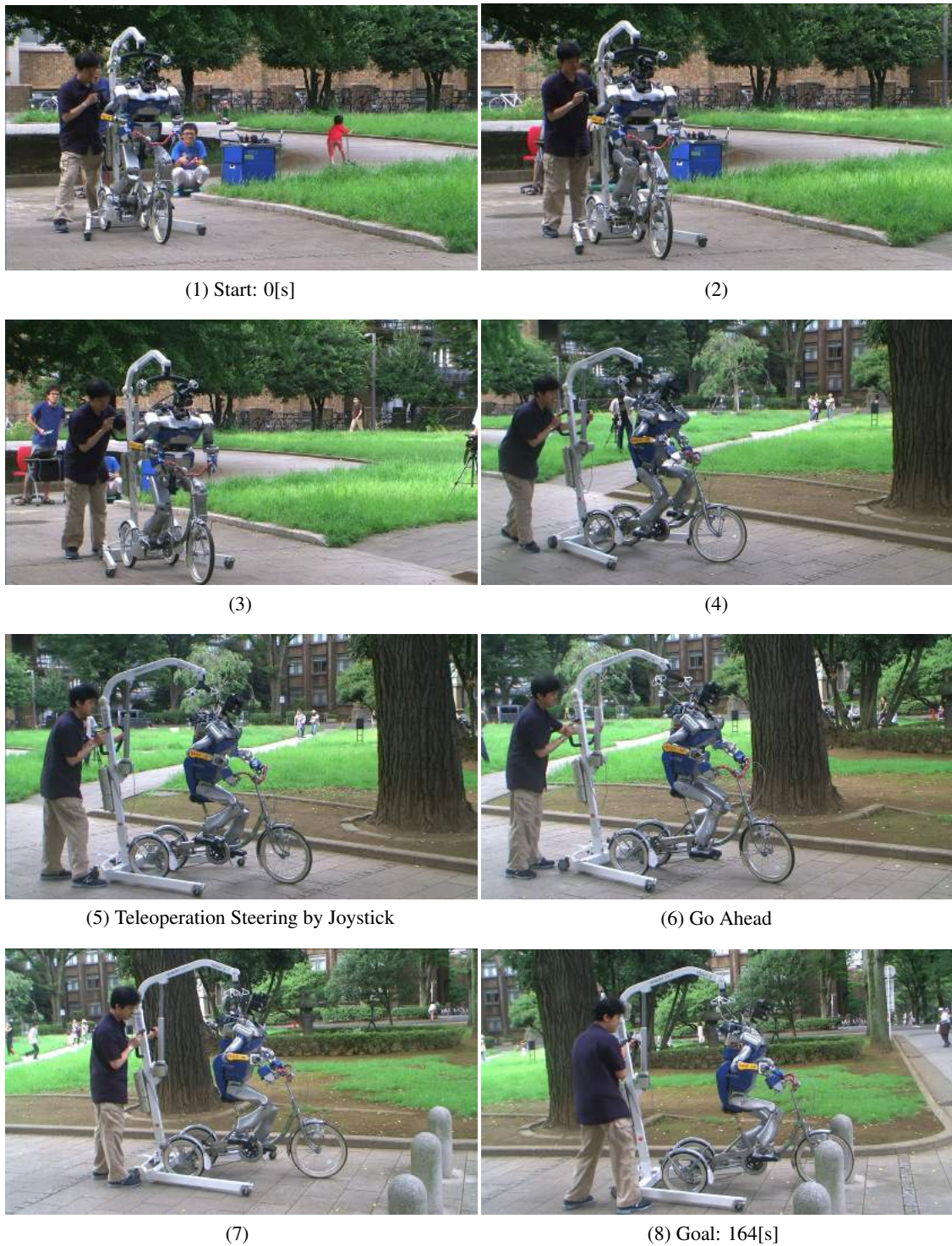


図 3.26: Steering and Pedaling by Joystick Controller Device in the Open Air (Front View)

た後にクランクを漕ぎ進める運転操作から、ハンドルを切ると同時にクランクも漕ぎ進めるような据え切りの発生しない運転操作へと改良)。

- 屋外環境における通信遅延の改善として、操縦用の計算機とロボット体内の計算機間のネットワークの通信を、無線による通信処理から有線LANによる通信処理へと変更。

上記に挙げた改良点を導入した屋外環境での等身大ヒューマノイドによる三輪車の運転操作行動の再実験結果のキャプチャを図3.27, 図3.28に示す。前回の実験で観察された走行の遅延性を低減させたことで、27.5[m]の屋外路面環境を図3.29のようにハンドル操舵も行いながら約1分12秒で走破した。この結果は、前回得られた平均の走行速度 7.8×10^{-2} [m/s] (= 12.8[m] / 164[s]) から大きく速度改善されていることが分かる。さらに本実験では、上述の改良点で示した能動機能と受動機能の間の目標接触力の差を大きくすることでクランクを介した内力の蓄積を分散させ、これにより重心まわりの外乱モーメントの発生による転倒要因を抑えている。この対策を加えることで、図3.27, 図3.28の実験結果で示したように等身大ヒューマノイドロボットの転倒防止のための補助リフトを外した、人間のサポート介入のない運転操作行動を実現している。

3.5.5 腱駆動ヒューマノイドによる運転操作

ヒューマノイドによる搭乗型機器の運転操作における別の観点からの実践的なアプローチとして、異なる身体構成をもつヒューマノイドに対しても提案した探索指標に基づく制御法の有用性を示していく。本項では、一般的な関節軸駆動型のヒューマノイドロボット HRP2-JSK [110] と異なる駆動構成を有するヒューマノイドとして、腱駆動ヒューマノイドロボット“腱志郎”[138][139]によるペダリング操作を実現した。著者の発表文献[126]を参考に以下で実現法を論じる。

腱駆動ヒューマノイドの三輪車操作における実験構成

腱駆動ヒューマノイドロボット“腱志郎”[138][139]を三輪車に搭乗させた構成を図3.30に示す。“腱志郎”は、重量が約51.9[kg]の筋骨格腱駆動ヒューマノイドロボットであり具体的な仕様・構成をA.2.2項に示す。“腱志郎”の腰リンクは三輪車に固定させた環境として双脚のペダリング操作による走行実験を行う。ロボットの脚の各リンク長(大腿リンク長: L_1 , 下腿リンク長: L_2 , 足リンク長: L_3), 及び固定点である股関節点 (Fixed Point) と同様に固定点と

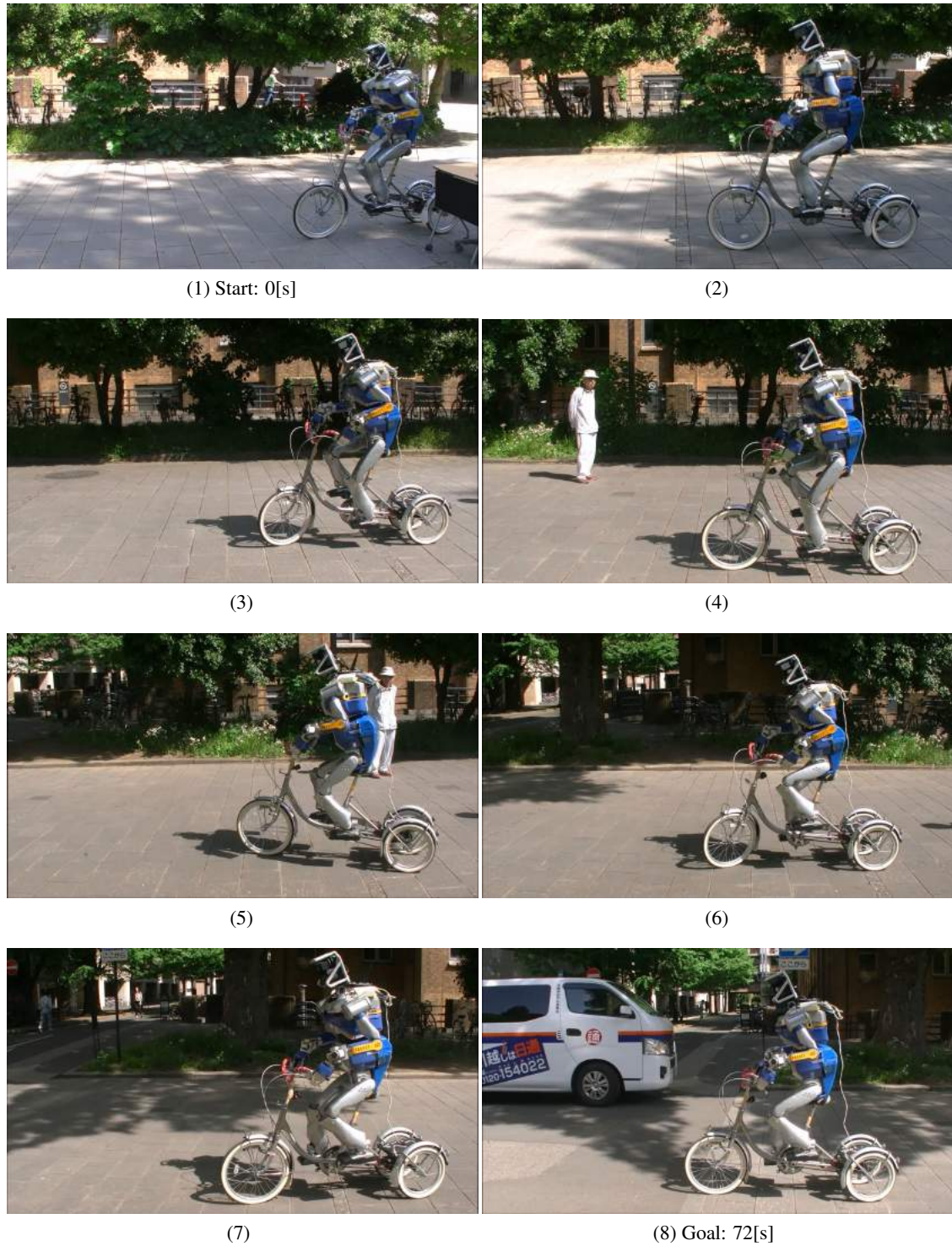


図 3.27: Improved Steering and Pedaling in the Open Air (Side View)

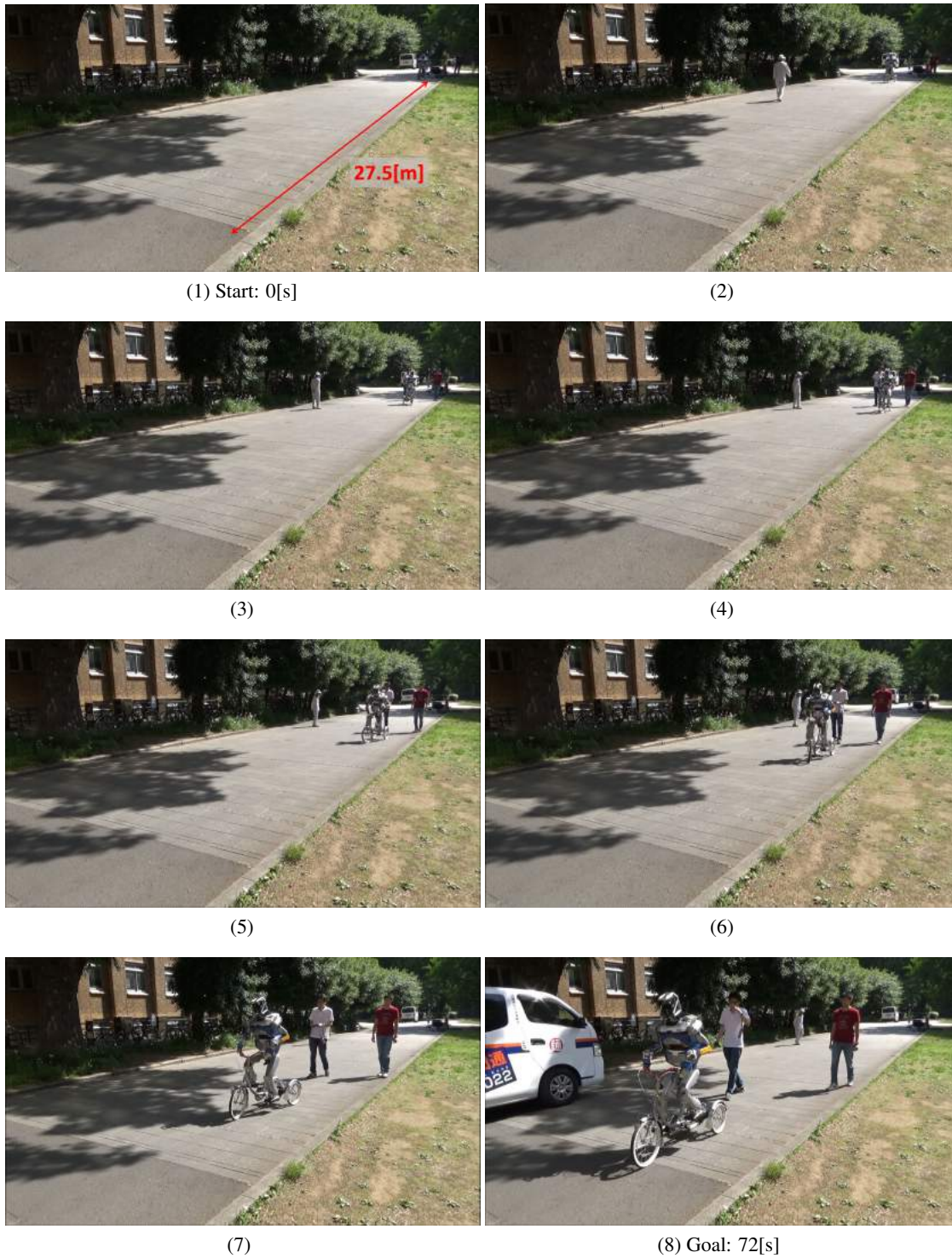


図 3.28: Improved Steering and Pedaling in the Open Air (Front View)



図 3.29: Foot View from Camera Attached to Tricycle

なるクランク中心 (Crank Center) までの x 成分, z 成分の距離 D_{cx} , D_{cz} , 足首関節点からペダル接触点までの距離 D_{ap} , クランク径 r はそれぞれ図 3.31 のように与えられる.

また, 本実験におけるコントローラの制御周期は 16[ms] とし, オンライン探索指標による軌道修正のために使用する足部搭載の CAN 通信規格の 6 軸力センサ (Istituto Italiano di Tecnologia 製) の分解能は, 並進 3 軸力について 0.25[N] の仕様である.

等身大ヒューマノイドロボット HRP2-JSK [110] による運転操作行動実験と同様に, 本実験においても 3.4 節で提案した探索指標に基づく能動機能と受動機能のマニピュレータ軌道の制御法がベースとなる. 式 3.7 で示したロボット足先の位置 \mathbf{p}_i の設定として, 筋骨格ヒューマノイド“臆志郎”の足部モデル [138] における \mathbf{p}_i の位置座標を図 3.32 に示す. 以降の議論では座標系もこれに従う.

能動機能を適用した脚とペダルの接触に関する拘束条件

筋骨格腱駆動ヒューマノイド“臆志郎”は身体構成の複雑性からインピーダンス制御 [48][49] やアドミッタンス制御 [50] をはじめとする力制御の適用・実装は発展途上の段階にある. そのため, 足裏とペダル間の接触力を十分に維持するための力制御法を用いた足-ペダルの離脱解消の方策に代わる特有の制御戦略が要求される. 本研究ではこの問題に対して, 3.4 節で述べたオフライン探索指標に基づく能動機能の軌道設計法に加えて, ペダルとの接触に関する拘束条件に基づいた足首関節のピッチ姿勢角の制御法を新たに提案する. 以下に導入した拘束条件を示す.

$$|(\mathbf{p}_{i+1} - \mathbf{p}_i) \cdot \mathbf{u}_i| > D_{th}, \quad i = 0, 1, \dots, T-2, T-1 \quad (3.21)$$

ここで \mathbf{u}_i は足裏面に対する単位法線ベクトル, D_{th} は閾値である. 不等式 3.21 の左辺は, 遷移ベクトル $\mathbf{p}_{i+1} - \mathbf{p}_i$ と単位法線ベクトル \mathbf{u}_i のなす角を Φ としたとき, 以下のように展開さ

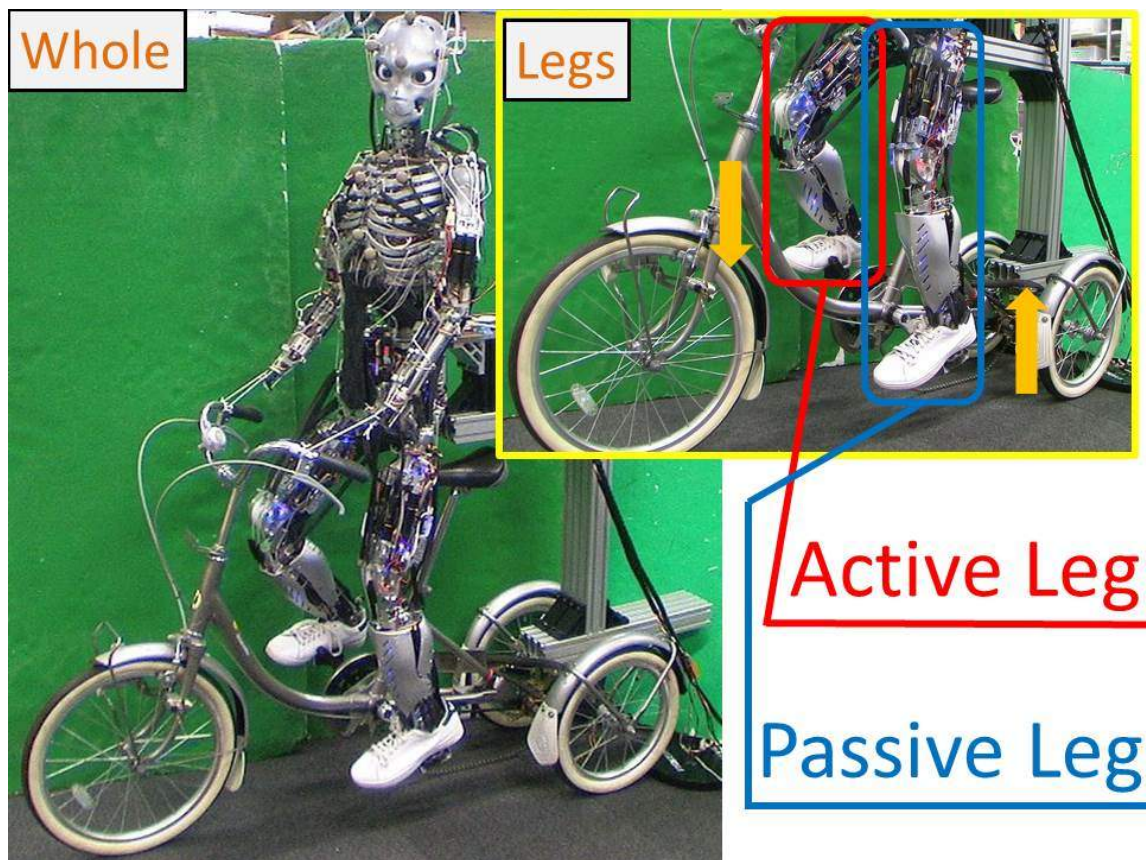


図 3.30: Applying Active/Passive Functions Based on Offline Gradient and Online Gradient to Tendon-Driven Humanoid [126]

れる.

$$\begin{aligned}
 |(\mathbf{p}_{i+1} - \mathbf{p}_i) \cdot \mathbf{u}_i| &= \|\mathbf{p}_{i+1} - \mathbf{p}_i\| \|\mathbf{u}_i\| \cos \Phi \\
 &= r \sqrt{2(1 - \cos \frac{\pi}{T})} |\cos \Phi|
 \end{aligned} \tag{3.22}$$

オフライン探索指標を用いた能動機能の作用時に足がペダルから滑り落ちないための許容範囲を $\Phi > 5\pi/6$ とするとき、閾値 D_{th} は以下の値となる.

$$D_{th} = 0.1 \sqrt{2(1 - \cos \frac{\pi}{18})} \left| \cos \frac{5\pi}{6} \right| \approx 0.015 \tag{3.23}$$

ここではクランク半径 $r = 0.1$ [m], 分割サンプリング数 $T = 18$ とした. 条件式 3.21 を満たすような \mathbf{u}_i を基に, 図 3.12 で設計した軌道への足先の追従を保ちつつ各時刻 i での足首姿勢

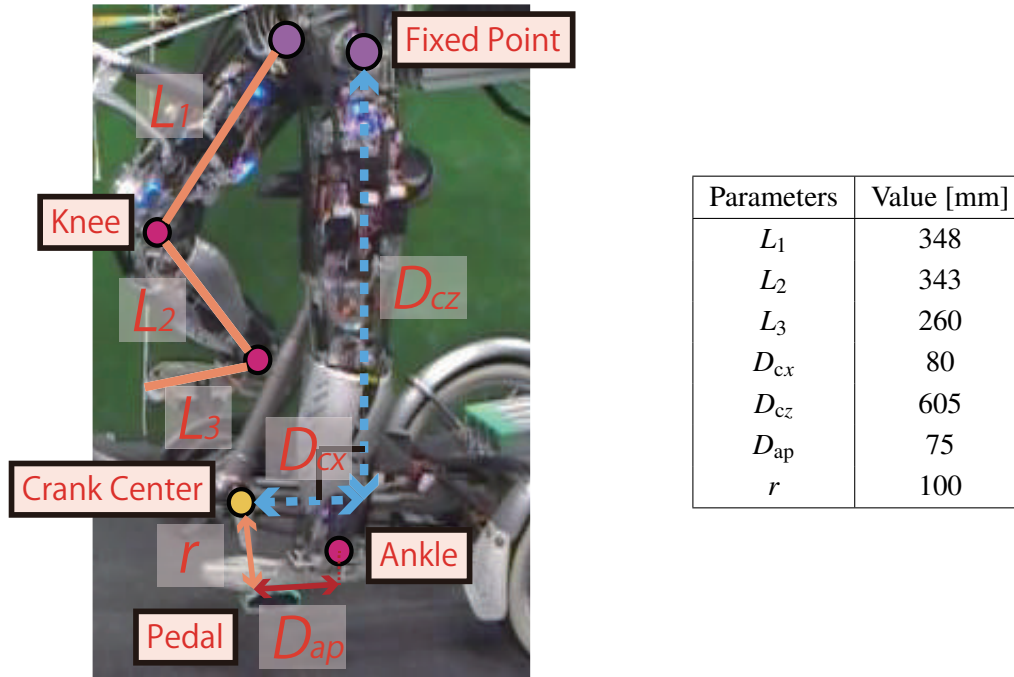


図 3.31: Experimental Parameters of Humanoid Leg and Tricycle [126]

角を制御する (図 3.33).

腱駆動ヒューマノイドにおけるオンライン探索指標

腱駆動ヒューマノイドによるペダリング操作においても、受動機能側の脚はオンライン探索指標を利用したペダルの接触検知による軌道修正法を適用する。これにより、受動機能側の脚はオンラインでペダルを回避しながら引き脚として機能することが可能である。時刻 i での受動機能側の脚のエンドエフェクタの位置 $\mathbf{q}_i = (x_{qi} \ y_{qi} \ z_{qi})^T$ 、および足が運動中のペダルから受ける外力変化 $\Delta \mathbf{f}_i = (\Delta f_{xi} \ \Delta f_{yi} \ \Delta f_{zi})^T$ (6 軸力センサにより取得) から、脚部エンドエフェクタの目標位置 $\mathbf{q}_{i+1} = (x_{qi+1} \ y_{qi+1} \ z_{qi+1})^T$ を更新し、 $\mathbf{q}_i \rightarrow \mathbf{q}_{i+1}$ に遷移をおこなう。このとき、更新則を以下の式で定義する。

$$\begin{aligned} \mathbf{q}_{i+1} &= \mathbf{q}_i + K_n \Delta \mathbf{f}_i \\ (&= \mathbf{q}_i + K (\Delta \mathbf{f}_i) \widehat{\Delta \mathbf{f}_i} \end{aligned} \quad (3.24)$$

ゲイン K_n はロボット足部搭載の 6 軸力センサより計測できる 3 軸力 $\Delta \mathbf{F}_s = (\Delta F_{xs} \ \Delta F_{ys} \ \Delta F_{zs})^T$ の変化量 (計測感度) $\|\Delta \mathbf{F}_s\|$ に対する受動機能側の脚の足先位置の所望の変位量パラメータ Δq_n

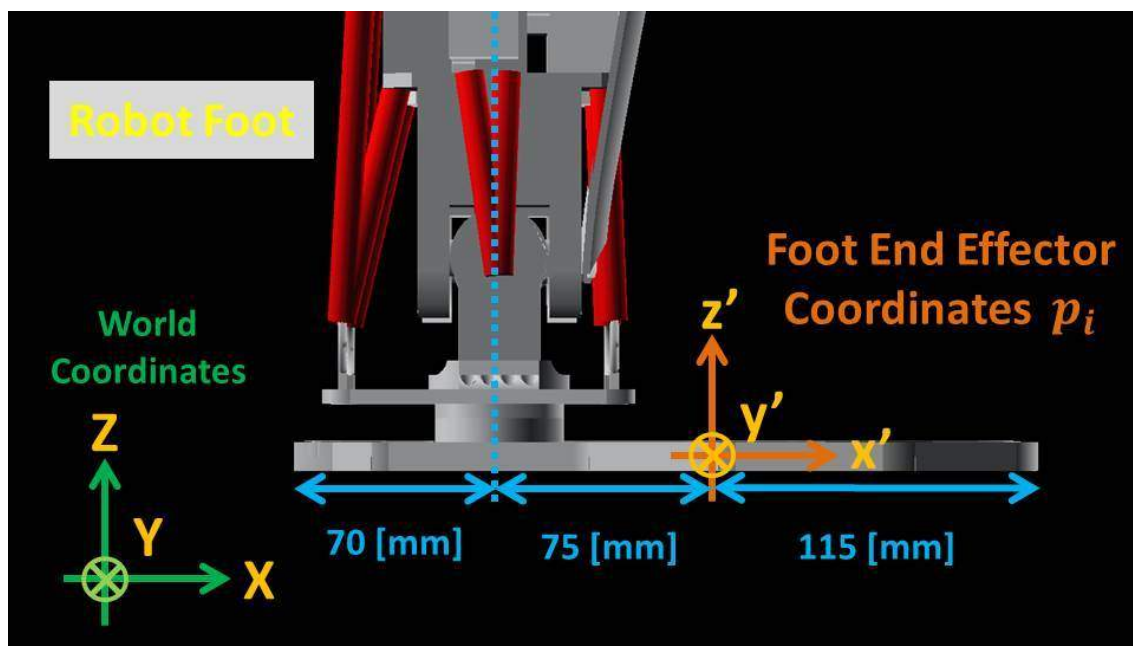


図 3.32: End-Effector Coordinates p_i on Humanoid Foot [126]

で定義される公称定ゲインであり,

$$K_n = \frac{\Delta q_n}{\|\Delta F_s\|} \quad (3.25)$$

と定める.

K_n は使用する 6 軸力センサおよび設定パラメータ Δq_n の情報により決まる一定量の比例ゲインとして, 本実験においては $K_n = 2[\text{mm}/\text{N}]$ と設ける. K_n は受動機能側の脚の足先エンドエフェクタの位置における剛性パラメータを表す.

また, 式 (3.24) は外力変化 Δf_i の単位方向ベクトル $\widehat{\Delta f}_i (= \Delta f_i / \|\Delta f_i\|)$ を分離することで, 各制御周期での 3 軸力に基づく更新ゲイン $K(\Delta f_i)$ を算出することができる. $K(\Delta f_i)$ は以下で表される.

$$K(\Delta f_i) = K_n \|\Delta f_i\| \quad (3.26)$$

以上より, ペダルから受ける外力変化のオンライン探索指標 $\widehat{\Delta f}_i$ に対する受動機能側の脚の移動量を $K(\Delta f_i)$ として動作更新が逐次的に可能となる. 式 3.24 において, 3.4.3 項で述べたオンライン探索指標による受動機能の定式化を参考にすると $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ 行列は

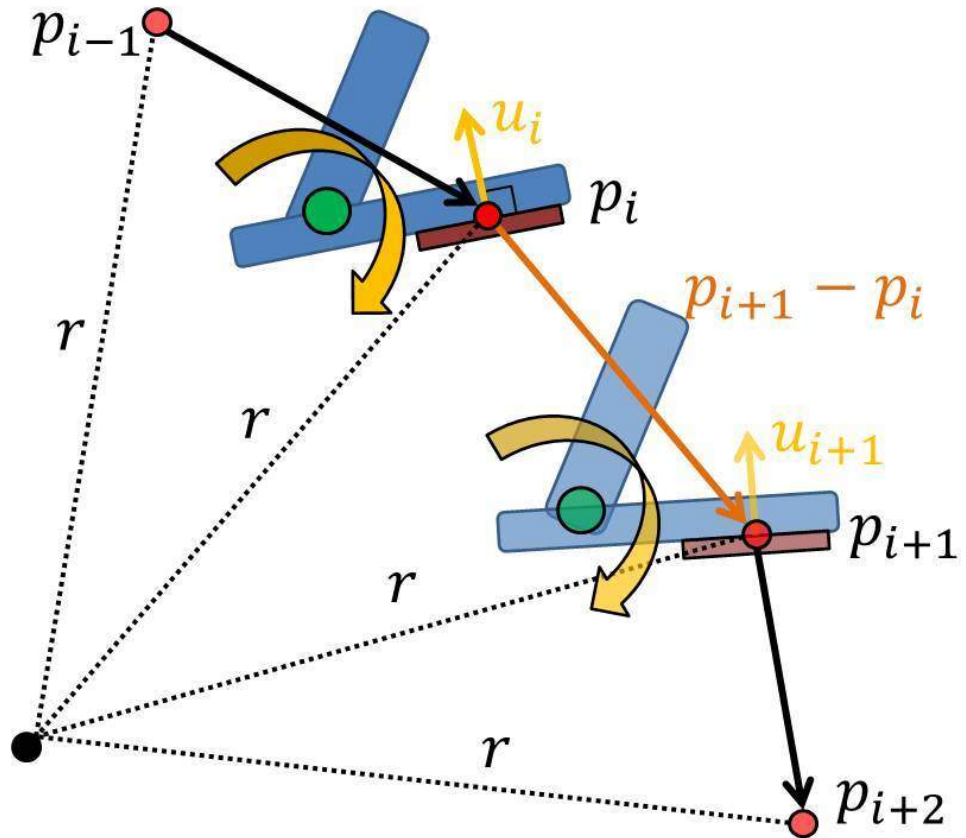


図 3.33: Ankle Posture Strategy of Active Leg [126]

式 3.17 と同様に表現できることから，オンライン探索指標および更新される軌道，一般化更新則の重み A は以下の対応となる．

表 3.5: 腱駆動ヒューマノイドにおけるオンライン探索指標の対応

オンライン探索指標	更新される軌道	一般化更新則の重み A
$\widehat{\Delta f}_i$	q_i	$K(\Delta f_i)$

腱駆動ヒューマノイドによるペダリング操作実験

オフライン探索指標に基づく能動機能，オンライン探索指標に基づく受動機能，の各機能の交互協調性を腱駆動ヒューマノイドの脚部に適用した，クランク 1 回転分のペダリング操

作の実証実験を行った。

実験結果 オフライン探索指標に基づく能動機能を適用した脚, およびオンライン探索指標に基づく受動機能を適用した脚のそれぞれの制御方策により, 最初に半回転のペダリング操作が実現された (図 3.34 -1,2,3 のフェーズ, 遷移移動時間:0[s]~約 5[s]). このとき, 能動機能を右脚, 受動機能を左脚に適用して走行を開始した. その後, 探索指標の切替則 (式 3.19) に基づき両脚の能動機能と受動機能を交換 (図 3.34 -3 から図 3.34 -4, 遷移移動時間:約 5[s]~約 6[s]) し, 再び半回転操作 (図 3.34 -4,5,6 のフェーズ, 遷移移動時間:約 6[s]~約 14[s]) することで1回転のペダリング操作を実現した.

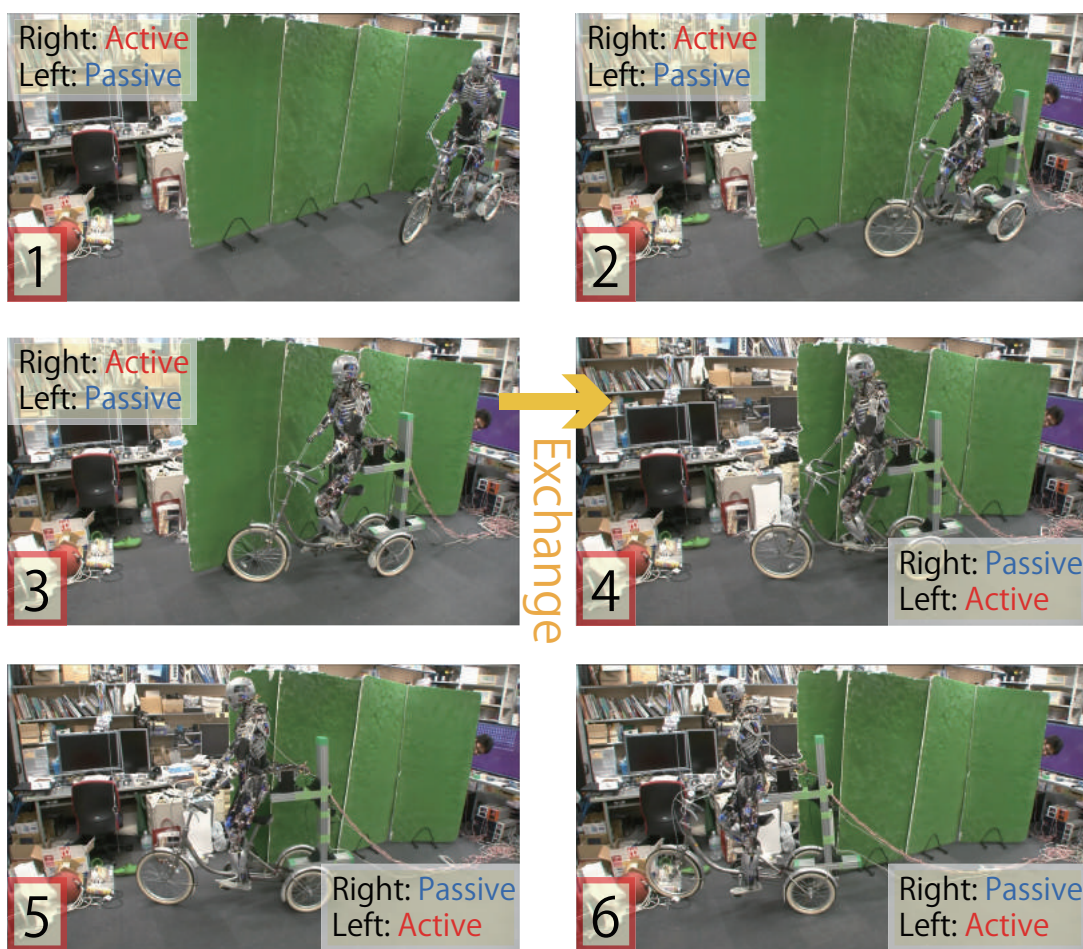


図 3.34: Snapshots of Tricycle Pedaling Experiment by Cooperative Manipulation of Active Leg and Passive Leg [126]

考察 腱駆動ヒューマノイドにおける能動機能と受動機能の協調作用について考察する。能動機能と受動機能を協調したペダリング操作時において、能動機能の脚および受動機能の脚が各々ペダルから受ける垂直反力 F_z の変化をクランク回し1回転分について計測した(図 3.35)。約 5~6[s] の能動機能と受動機能の切替期を境に、各脚の能動期/受動期が交換されている。

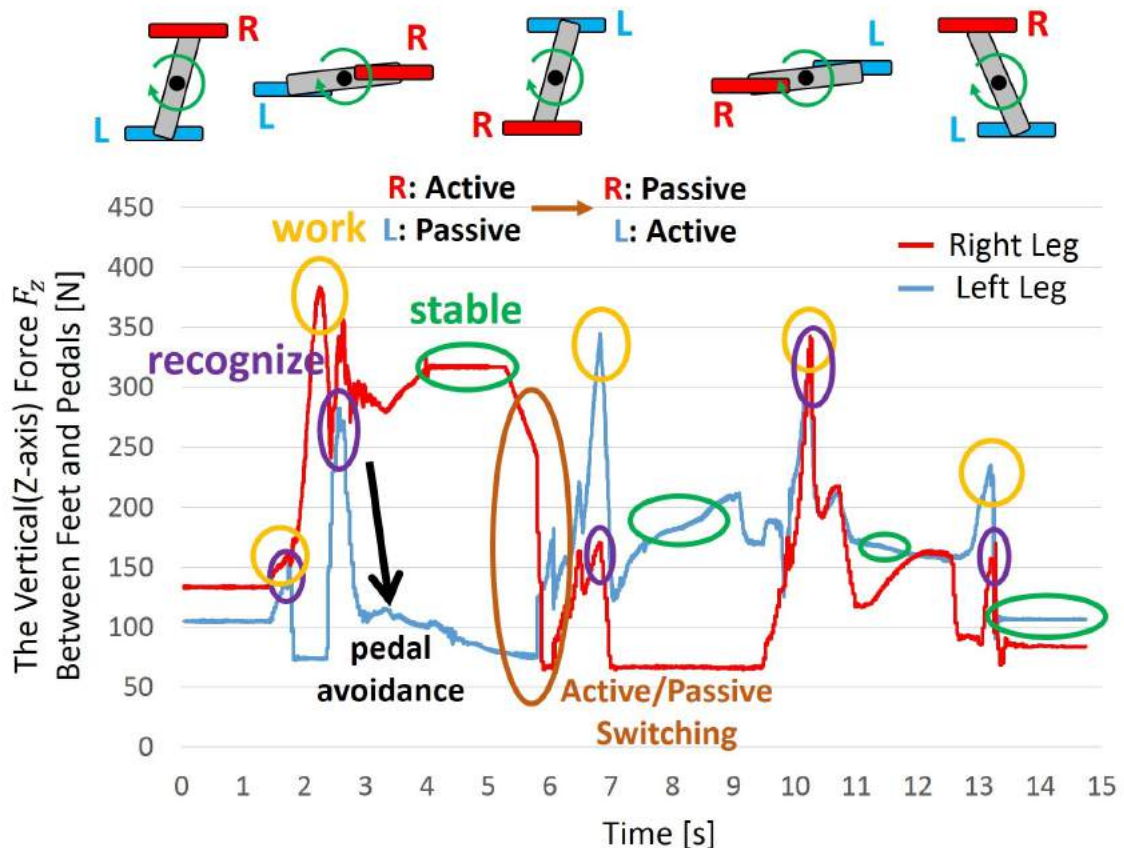


図 3.35: Changes of the Vertical (Z-axis) Force F_z [N] between Feet and Pedals [126]

図 3.35 より次の点が考察される。

- 各時刻 (2,3,7,10,13[s] 付近) において、能動機能側の脚が生成する踏み込み力 (図 3.35 - 黄丸囲) がクランクを介して受動機能側の脚に伝達し、内力負荷が高まったのちに、力検出 (図 3.35 - 紫丸囲) によるオンライン探索指標に基づく制御法 (式 3.24) に従って、ペダルの回避動作が実現された。

- ピークとなる力検知後、受動機能側の脚のペダル回避動作によりクランクの回転障害を緩和し、次の受動機能側の脚の力検知まで能動機能側の脚が安定した力でペダルの踏み込みを実現した(図 3.35 -緑丸囲).
- 5~6[s]での能動機能側の脚と受動機能側の脚の切替操作の前後において、能動機能の脚の踏み込み力が右脚(4~5[s]で約310[N])と左脚(8~9[s]で約170[N])で異なっている。これは切替操作による足裏とペダルの接触位置のズレ、また柔軟な腱駆動ロボット特有の腱ワイヤの伸びによる姿勢変化・接触力方向(鉛直 z 成分)からのズレの影響が考察される。

このようにロボット固有のハードウェア特性の影響誤差に対しても、クランクの環境を介して適切に受動機能側の脚を位置更新させることで実環境に応じた柔軟な制御を可能とした。

- 切替操作後における10[s]付近では受動機能側である右脚がペダルから大きな反力を受けている。このときのクランク角度は図 3.35 -上図の通り水平に近く、以前の時点と異なり能動機能側(左脚)の踏み込み力が鉛直成分について高効率で受動機能側(右脚)へと伝達される。

受動機能側(右脚)は大きなペダル反力を受けるも、この力(オンライン探索指標)に比例した位置の軌道修正(式 3.24)を実行することでペダルから回避する負荷軽減動作を実現していることが11[s]付近で観察される。この現象は切替操作前の2~4[s]においても同様に観察できる。

以上の観察事項より、オフライン探索指標に基づく能動機能とオンライン探索指標に基づく受動機能はクランクの環境を媒介として、交互の協調作用を担っていることが実証された。

3.6 本章のまとめ

本章では、ヒューマノイドにおける搭乗型機器の操作に関して、ヒューマノイドが搭乗型機器を操ることの課題を述べ、それに対する行動戦略について述べた。計画した行動戦略を実現するためには、ヒューマノイドがもつ複数のマニピュレータ（双腕と双脚）を協調させた軌道制御法が必須であり、その制御方策を本章で提案した。本章では、協調する2つのマニピュレータに対して2つの制御機能を与えており、その制御を実現するために第2章で提案した探索指標の考えを利用している。以下にそれぞれの探索指標に基づく各機能の役割を整理する。

1. オフライン探索指標ベースの能動機能:

操作対象物（ハンドルやクランク）に対して所望の位置や力を与えるマニピュレータの軌道計画法

2. オンライン探索指標ベースの受動機能:

操作対象物（ハンドルやクランク）の運動に応じてマニピュレータの軌道に逐次修正を加える軌道修正法

上記のような2つの探索指標の概念を協調させ切り替えることで、双腕によるステアリング操作と双脚によるペダリング操作の制御課題に共通の解を与えた。さらに、提案した探索指標に基づく制御方策の有用性を示すため、室内環境および屋外環境の2種類の走行環境、また関節軸駆動型ヒューマノイドロボットおよび腱駆動ヒューマノイドロボットの異なる身体構成をもつ2種類のヒューマノイドに対しても提案手法が活用できることを実機実験により実証した。

第4章

運転操作における 搭乗姿勢と外界環境に対する軌道修正

4.1 はじめに

本章では、第3章で提案した搭乗型機器の操作に対する軌道修正戦略を補助するヒューマノイド自己の搭乗姿勢の安定化、およびヒューマノイドと搭乗型機器の外界に存在する外乱環境に対する認識行動について、軌道修正の側面からのアプローチを提案する。

4.2節では、ヒューマノイドにおける搭乗姿勢の安定化と外界環境の認識に関して論点を整理し、それぞれが解くべき課題について述べる。4.3節では、搭乗姿勢の安定化を目的とした軌道修正戦略を示す。搭乗姿勢の安定化においては、搭乗型機器上でヒューマノイドが受ける接触拘束を解除した離脱行動に基づく“割り込み処理型座り直し制御”の方策と、運転操作行動とマルチスレッドに実行可能な“実時間並行処理型姿勢制御”，の2つのアプローチを提案し、互いの特徴について論じる。4.4節では、ヒューマノイドと搭乗型機器からなる系の外に存在する外界環境について、視覚認識に基づいた軌道修正により解かれる大型障害環境と、力覚検知に基づいた軌道修正により解かれる小型障害環境、の2つに外乱因子を切り分けて議論する。

4.2 ヒューマノイドにおける搭乗姿勢の安定化と外界環境の認識

ヒューマノイドの運転操作において「搭乗姿勢の安定化」と「外界環境の認識」は一見すると異なる2つの概念であるが、両者はともにどの視点からみたヒューマノイドの行動戦略か、という点で共通に捉えることができる。すなわち、図4.1に示すようにヒューマノイド系 Σ_{humanoid} が行動の主体となり、どの環境（搭乗型機器系 Σ_{vehicle} ，外界環境系 Σ_{outside} ）からみた情報を基に動作計画や動作生成を実行するかが論点となる。

図4.1に基づいて「搭乗姿勢の安定化」と「外界環境の認識」を理解すると、次のような軌道修正戦略と考えることができる。

搭乗姿勢の安定化:

Σ_{vehicle} に対する Σ_{humanoid} の軌道修正

外界環境の認識:

Σ_{outside} に対する Σ_{humanoid} と Σ_{vehicle} の軌道修正

搭乗姿勢の安定化の概念と比較して、外界環境の認識においてはヒューマノイド系 Σ_{humanoid} に加えて搭乗型機器系 Σ_{vehicle} への操縦軌道の修正も考慮する必要がある。以下で具体的な

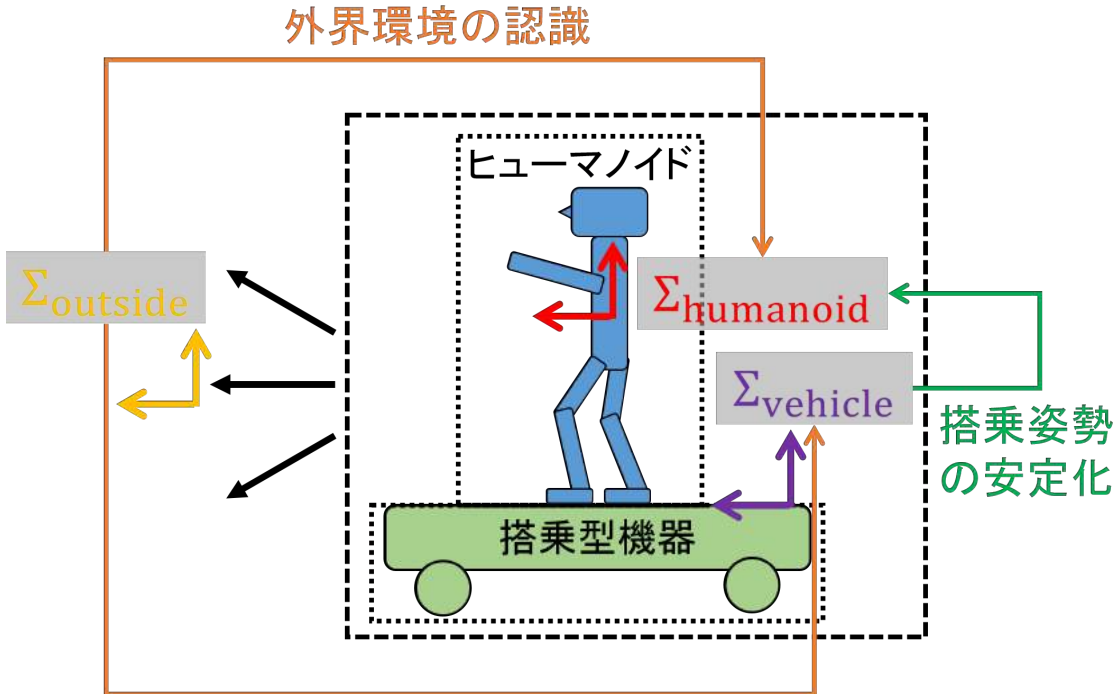


図 4.1: 3つの系（ヒューマノイド系 Σ_{humanoid} ，搭乗型機器系 Σ_{vehicle} ，外界環境系 Σ_{outside} ）からみた「搭乗姿勢の安定化」と「外界環境の認識」の関係

例も与えながら，ヒューマノイドにおける搭乗姿勢の安定化と外界環境の認識で要求される軌道修正課題および戦略を述べる．

4.2.1 軌道修正が解く搭乗姿勢の安定化に対する課題

ヒューマノイドによる搭乗型機器の運転操作では，運転するための搭乗型機器への操作・マニピュレーションに加えて，運転操作時のヒューマノイド自身の搭乗状態の姿勢（搭乗姿勢）の安定化を図る制御方策が必要不可欠である．この要求課題は二輪車等の搭乗型機器自体に不安定性を有する課題に限らず，三輪車のような搭乗型機器自体は安定的であるが搭乗するヒューマノイドの不安定性により結果的に不安定となる課題でも問題となる．

一般に歩容生成をはじめとするヒューマノイドの姿勢のバランス維持の提案手法として，ZMP（Zero Moment Point）[60][142]の概念がこれまで幅広く活用されてきた[143][144]．しかしながら，ZMPの安定化指標の考えは接触面が同一平面上に存在するときのみ有効で

あった。そこで、Haradaら [145][146] は三次元空間における物体との接触などの、接触面が同一平面上にないときのバランスの指標として、ZMPを拡張した一般化ZMPを提案しヒューマノイドによる安定的な脚腕協調動作を実現した。

搭乗型機器の操作の例として、ヒューマノイドによる三輪車操作においても一般化ZMPの考えは有用であるといえるが、支持集合の凸包を得るにあたり、手とハンドル、足裏とペダル、臀部とサドルなど各々に対する接触支持領域を厳密に定義する必要がある。特に臀部とサドルの接触においては接触領域の把握が困難であり、ヒューマノイド固有の臀部のデザインや搭乗する三輪車のサドル形状に対しても変化的となる。また、Haradaらによる脚腕協調動作から発展させた、臀部-サドルの接触支持を考慮した新たな拡張アプローチが要求される。

本研究では、三輪車を例にヒューマノイドと三輪車の接触領域のうち、圧力が最も大きくかかる接触域として臀部-サドル間を考え、不安定性の要因を

- 臀部-サドルの接触領域における領域面積の変動
- 臀部-サドルの接触領域からの重心投影点の外れ

が支配的であると切り分けて、これらを修復する軌道修正戦略を提案する。上記2つの転倒要因を背景に、具体的な軌道修正戦略を挙げると以下のような方策となる。

搭乗姿勢の安定化を目的とした軌道修正

- ヒューマノイド臀部とサドル間の着座面の補正
- ヒューマノイドの重心投影点のサドル上への修復

さらに、本研究ではこれらの軌道修正の実行処理について、搭乗型機器の操作行動と並行した実時間処理型の姿勢制御法と、操作行動中に割り込んで処理する割り込み処理型の座り直し動作に基づく安定化手法、の2通りの軌道修正法を提案する。

4.2.2 軌道修正が解く外界環境の認識に対する課題

運転操作行動中はヒューマノイドの自己姿勢のリカバリーに留まらず、ヒューマノイドと搭乗型機器から構成される系以外の外界の環境 Σ_{outside} を把握し、それに応じた行動の軌道修正を図る必要がある。外界の環境は多種多様であるが、本研究では運転操作の妨げとなる障害の外乱要因に着目する。

運転操作中の外乱要因を大きく分類すると、

$$\text{外界に存在する外乱} \begin{cases} \text{大型障害物による外乱} \\ \text{小型障害物による外乱} \end{cases}$$

に分別できる。それぞれの障害物の環境の例を表 4.1 に示す。

Large Obstacles	Small Obstacles
Wall	Step
Partition	Grass
People	Debris
Other Robots	Slope

表 4.1: Examples of Large Obstacles and Small Obstacles

大型障害物による外乱は、単純突破が困難な障害環境であり回避する行動戦略が望まれる。一方、小型障害物による外乱は、ロコモーションにおける足元（搭乗型機器では路面に相当）に存在する僅かな低い段差などの、ロボットの視覚を利用して認識が困難であり回避行動よりも突破行動が推奨される障害物環境と定義する（図 4.2）。

これらの障害物環境における行動戦略では、異なる外界環境の認識アプローチが要求される。大型障害物を含む環境下においては、ロボットの目線の行き届く範囲内に障害物を環境認識点群などを通じて視野に収めることが可能であり、ロボット頭部に搭載されたステレオカメラなどを用いた視覚認識行動が有効な手段であると言えよう。視野の行き渡りやロボットのビジョンを利用した認識行動が困難な床面の障害物・地形に対しては、搭乗型機器を介した力学的な認識法が新たに要求される。三輪車操作の例においては、ヒューマノイド脚部によるクランクのペダリング漕力の変化・変動の把握、すなわち漕力の認識が重要となる。

以上で論じた方法論を整理すると、次のような障害物環境に応じた軌道修正戦略が外界環境の認識において解くべき方策となる。

- 「大型障害物」に対する外界環境認識行動
⇒ “視覚情報” のナビゲーションに基づいた運転操作の軌道修正
- 「小型障害物」に対する外界環境認識行動
⇒ “力覚情報” のナビゲーションに基づいた運転操作の軌道修正



図 4.2: Various Obstacles (Classification to Large or Small)

Large Obstacles: Partition, Other Robots, Miscellaneous Objects, Small Obstacles: Grass, Step, Debris

4.3 搭乗姿勢の安定化に対する軌道修正法

本節では、ヒューマノイドにおける搭乗型機器の運転操作行動中に発生する姿勢のズレ・傾きを探索変数にとるオンライン探索指標（オンライン制御指標）を利用した軌道修正法を提案する。搭乗姿勢の安定化を図る修復方策として、本研究ではメインとなる操作行動タスクに対して、並行した処理系もしくは割り込む処理系の2通りの実行処理方法を提案し、それぞれの特徴について論じる。本論文においては各処理系の制御方式を以下のように呼ぶものとする。

- 割り込み処理型座り直し制御
- 実時間並行処理型姿勢制御

特に、割り込み処理型座り直し制御で実行されるヒューマノイドの座り直し動作では、3.4節で提案したマニピュレータ軌道の能動機能と受動機能の制御法を共通利用した行動戦略を述べる。

4.3.1 運転操作時の着座姿勢のズレにおける問題点

ヒューマノイドによる三輪車操作の行動例では、ヒューマノイド身体と三輪車との接触拘束状態の変化による姿勢のズレが修復すべき課題となる。特に表 4.2 に示すように、足裏とペダルの接触拘束、臀部とサドルの接触拘束はユニラテラル拘束であり、非接触方向に対しての離脱が問題である。

表 4.2: Contact States (Bilateral or Unilateral) for Humanoid Body and Tricycle

	Bilateral Contact	Unilateral Contact
Right Hand - Right Handle Grip	○	×
Left Hand - Left Handle Grip	○	×
Right Foot - Right Crank Pedal	×	○
Left Foot - Left Crank Pedal	×	○
Hip - Saddle	×	○

本研究では、双脚によるペダリング操作に関して足裏とペダルの鉛直方向の接触における拘束力を一定に保つ制御法を提案しているため、足-ペダル間の接触力は摩擦円錐内に抑えることが可能である。一方で、臀部とサドルの接触においては

- ヒューマノイドのエンドエフェクタ以外での接触のため、接触力・摩擦力の制御が困難
- 臀部表面形状・サドル面のデザインにより接触領域の定義が曖昧

なため、ズレ・離脱の発生が懸念される。そのため、本項では主にヒューマノイド臀部とサドルの接触状態の変化により引き起こされる着座姿勢のズレに注目して議論する。

着座した座面のズレにより引き起こされる問題点として以下の影響が考えられる。

(問題 1) 姿勢の変化に伴う総重心の変動による転倒

(問題 2) 重心移動による脚腕にかかる負荷の増大

(問題 3) 各接触箇所(ハンドル把持点・ペダル接地点)へのズレの伝搬・蓄積

図 4.3 左図は座面ズレが発生している状況下での三輪車運転時のキャプチャであり、右図がそのときに監視したヒューマノイドが受ける負荷(手先・足先の接触力及び各関節にかかる負荷)のモニタである。手先・足先にかかる力は6軸力覚センサによる計測量として矢印で

表現される。関節にかかる負荷は関節位置追従エラーから推定した値に応じて円形の矢印で表されている。円の半径が関節負荷の絶対値に対応している。

図 4.3 右図を観察すると、左側に重心が移動しており、反時計まわりに転倒の可能性(問題 1)があがっていることがわかる。また、左手首関節の Roll 方向で特に負荷が大きくみられる。座面ズレにより重心が移動したことで、計画軌道から乖離し接触力の誤差が大きくなり、関節トルク負荷(問題 2)が大きくなったと考えられる。(問題 3)については後述の 7.4.1 項で検証を行う。

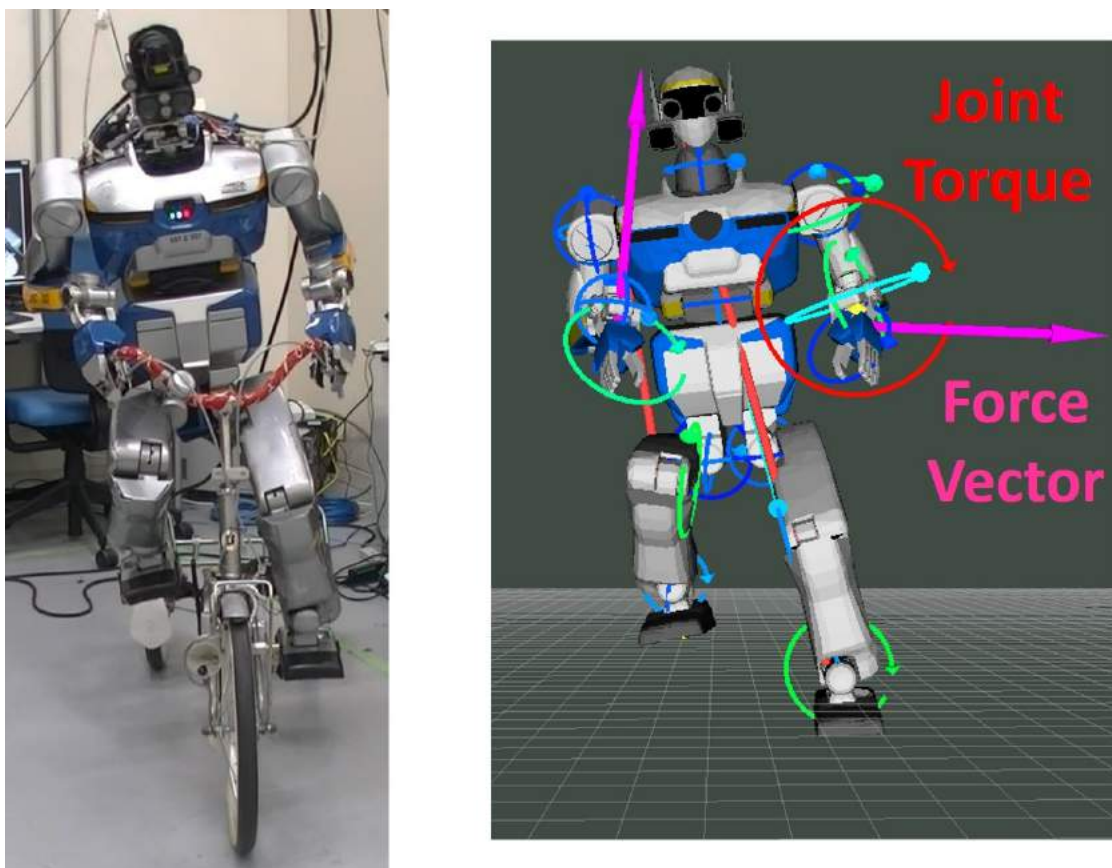


図 4.3: Snapshot of the Inclined Humanoid during Driving, and the Monitoring Viewer about Force Vector and Joint Torque [52]

これらの問題の要因となるヒューマノイド臀部のサドル座面に対するズレについて、大きく 2 つに分けて考察すると次のようになる。

臀部のサドル座面に対する並進ズレ

ヒューマノイド臀部とサドル面の接触ズレの要因として、並進的な座面ズレについてまず議論する。環境に固定されたサドルに対する臀部の並進ズレは図 4.4 -(i) に示す通り、前後方向 (x 方向) および横方向 (y 方向) において発生する。不安定性を呈する (問題 1) で挙げたように、ズレによる重心変動について考える。4.2.1 項で述べたように、安定性の条件として臀部とサドルの接触領域 $\mathbf{S}_{\text{hip}}^{\text{saddle}}$ に対して、重心投影点 $(x_G, y_G)^T$ は以下の条件を満たす必要がある。

$$\begin{pmatrix} x_G \\ y_G \end{pmatrix} \in \mathbf{S}_{\text{hip}}^{\text{saddle}} \quad (4.1)$$

重心投影点 $(x_G, y_G)^T$ は領域 $\mathbf{S}_{\text{hip}}^{\text{saddle}}$ 内に存在することが望まれる。重心投影点の x 方向, y 方向へのそれぞれの許容範囲を考えると、図 4.4 -(i) において x 方向はサドルノーズが存在する一方で、 y 方向は肉薄となっている。そのため、 y 方向への重心ズレが極めて支配的であることから、本研究では y 方向に対する並進ズレを修復するような動作方策に焦点を当てる。

臀部のサドル座面に対する回転ズレ

次に回転ズレについて議論する。図 4.4 -(ii) に示す通り、臀部の回転ズレは x まわり, y まわり, z まわりで発生するが、サドルのノーズにあたる x 方向への領域の伸びにより、臀部の y まわり (Pitch まわり) のモーメントはサドルからの反モーメントとして吸収が可能である。また、 z まわり (Yaw まわり) の臀部回転は、接触領域内に重心投影点が存在する場合 (安定的である場合)、投影点における自重に対する垂直抗力により十分な静止摩擦力が得られる。

以上より、サドル座面における臀部の回転ズレは x まわりが寄与し、特に y 方向への重心の並進ズレによる x まわりの自重モーメントが要因となる。

4.3.2 割り込み処理型座り直し制御

前項で議論した搭乗時の着座姿勢の y 方向の並進ズレ, x まわりの回転ズレを補正する目的の安定化制御法を本項で提案する。本項の内容は、筆者主著論文 [52] において詳しくまとめられており、この内容を基にしたオンライン探索指標 (オンライン制御指標) を用いたアプローチについて論じる。

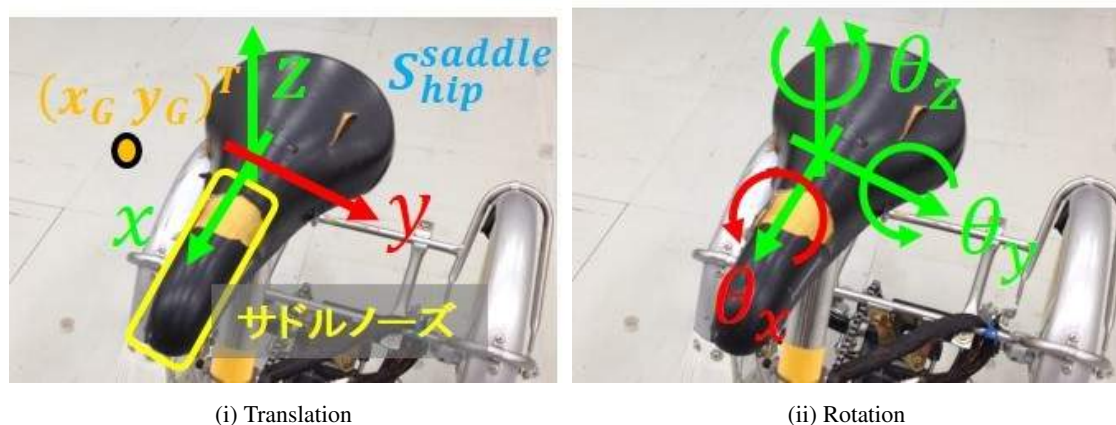


図 4.4: Translational and Rotational Contact Error between Humanoid Hip and Tricycle Saddle
For translational contact error, Y-axis element is dominant. For rotational contact error, X-axis element is dominant.

本項で述べる安定化制御法は、ヒューマノイドによる運転操作行動のタスク実行中において割り込み処理として働く切替処理型の実装行動となっている。運転操作行動中に割り込む動作生成は、ヒューマノイドが搭乗型機器の座面（サドル）から腰を浮かして身体の重心移動を行う座り直し動作に基づく処理方法であることから、この安定化制御法を「割り込み処理型座り直し制御」と定義する。

割り込み処理型座り直し制御では、HRP2-JSK [110] 体内に搭載されたジャイロセンサ・加速度センサに基づいた体幹姿勢推定を行うことにより、体幹の Roll 方向の傾きを把握し、その傾斜角に応じた座面ズレのフィードバック補正を実現させる。体幹のサドルに対する位置ズレと姿勢傾斜の両方を補償するため、体幹重心の移動により補正動作を達成させる。

着座姿勢のズレの検知

図 4.3 のキャプチャで示したように、搭乗状態にあるヒューマノイドが大きく傾斜した場合の着座姿勢のズレの自己診断系をはじめに構築する。y 方向に対する並進ズレの検知方法として、ロボットの自己姿勢の並進位置の獲得は自己位置推定法 [147][148][149] をはじめとして、様々な手法が提案されてきている一方で、本研究で目的とするような搭乗型機器を相対としたヒューマノイドの位置ズレ等の僅かな並進変動に対する計測手法はほとんど確立されていない。[150] のようにロボットに搭載した IMU から加速度項の積分により位置を推定する手法が提案されているが、三輪車に搭乗したヒューマノイドの場合、三輪車相対の位

置ズレは微小変動により計測誤差の伝搬が目立つ。

そこで、 y 方向に対する並進ズレが x まわりに対する回転ズレに寄与すると考え、 y 方向に対する並進ズレ、 x まわりに対する回転ズレを、1つのセンサインターフェースにより検知する手法を採用する。すなわち、 x まわりに対する回転ズレのセンシングを目的とする。本研究ではHRP2-JSK [110]体内に搭載されたIMU (Inertial Measurement Unit)により、体幹のRoll方向の傾きを x まわりの回転ズレとして把握する。

双脚の能動的踏み込みを利用した臀部の接触離脱

第3章で提案したハンドルのステアリング操作やクランクのペダリング操作では、能動機能と受動機能の1対1の組み合わせにより実現された。割り込み処理型座り直し制御では、ステアリングやペダリングなどの操作行動は実行せず、ヒューマノイド自身の姿勢補正を目的とした、双脚ともに能動機能を与える制御戦略を構築する。この戦略は座り直し動作の生成の根幹をなす行動戦略であり、座り直し動作におけるステアリング操作行動やペダリング操作行動との能動機能/受動機能の適用法の差異と立ち位置を明確化した図を図4.5に示す。

座面ズレの補正戦略として、ここでは臀部-サドル間の接触座面の一時解消(離脱)を図り、重心の位置補正を行う方策を採用する。これは接触面における摩擦力を軽減するためである。一般に臀部-サドル間に働く垂直抗力 $F_{\text{hip},z}$ は右足,左足,右手,左手にかかる垂直抗力 $F_{\text{rf},z}, F_{\text{lf},z}, F_{\text{rh},z}, F_{\text{lh},z}$ から図4.6-(I)の動力学モデルに従い、式4.2のように書き表される。ここで $F_{\text{hip},z}, F_{\text{rf},z}, F_{\text{lf},z}, F_{\text{rh},z}, F_{\text{lh},z}$ は目標力を設定している。また、このときのクランク角の姿勢は両足平が水平となるように $\theta_c = k\pi$ (k :整数)と定め、この状態を座り直し動作の実行時期とする。

$$F_{\text{hip},z} = M(\ddot{z}_g + g) - (F_{\text{rf},z} + F_{\text{lf},z} + F_{\text{rh},z} + F_{\text{lh},z}) \quad (4.2)$$

ここで M はヒューマノイドの総質量、 \ddot{z}_g は総重心点の鉛直上方(z 方向)の加速度、 g は重力加速度を表す。 $F_{\text{hip},z}$ を小さくして0に近づけるためには、 \ddot{z}_g を小さくする、もしくは $F_{\text{rf},z} + F_{\text{lf},z} + F_{\text{rh},z} + F_{\text{lh},z}$ の項を大きくする必要がある。腰の浮上に伴い \ddot{z}_g の増加は避けられないため、後者の値が大きくなるような方策を選択する。特に3.3.3項で述べたように、両足とペダルの接触力が大きく寄与することから、 $F_{\text{rf},z}, F_{\text{lf},z}$ の能動的な踏み込み増加により臀部のサドル離脱を図る。

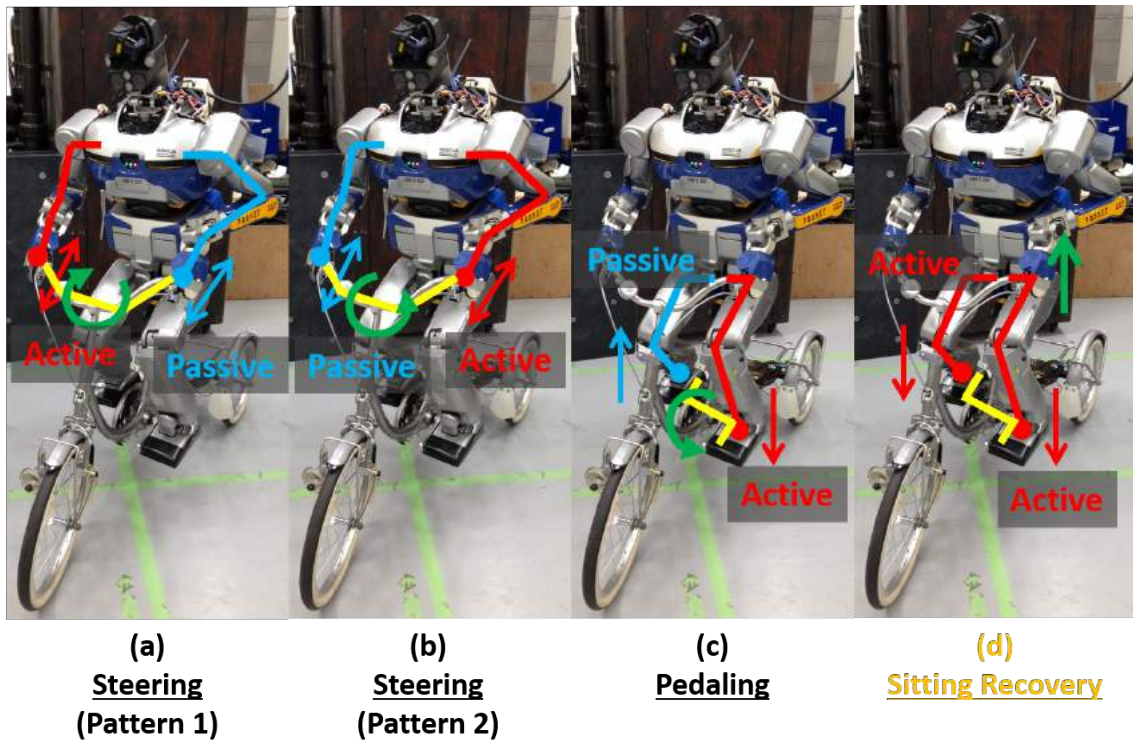


図 4.5: Each Active/Passive Functional Classification [52]
Sitting recovery motion (d) is based on the bipedal active standing.

体幹傾きフィードバックに基づくズレ補正動作

臀部の接触離脱後は左右への体幹移動が可能となる. 図 4.6 -(II) のようにヒューマノイドの体幹重心移動をおこない, 重心の y 方向へのズレ補正量 Δy_g を式 4.3 に基づき制御する.

$$\Delta y_g = L_{\text{hip} \rightarrow \text{COG}} \sin \theta_{\text{roll}} \quad (4.3)$$

$L_{\text{hip} \rightarrow \text{COG}}$ は臀部から重心までの距離, θ_{roll} は Roll 方向の姿勢傾き角である. ロボット体内に搭載された IMU から取得される θ_{roll} をズレ補正量 Δy_g にフィードバックすることで座面ズレの補正を実現する. ズレ補正動作後は双脚の両足-ペダル間の目標接触力を初めの $f^{\text{trg}}(t) = (0 \ 0 \ 150)^T$ [N] に戻し, サドルへの腰掛け動作を行う.

オンライン探索指標に基づく重心の軌道修正のアプローチ

式 4.3 で定めた重心のズレ補正量 Δy_g は, 重心の軌道に対する修正項として捉えることが

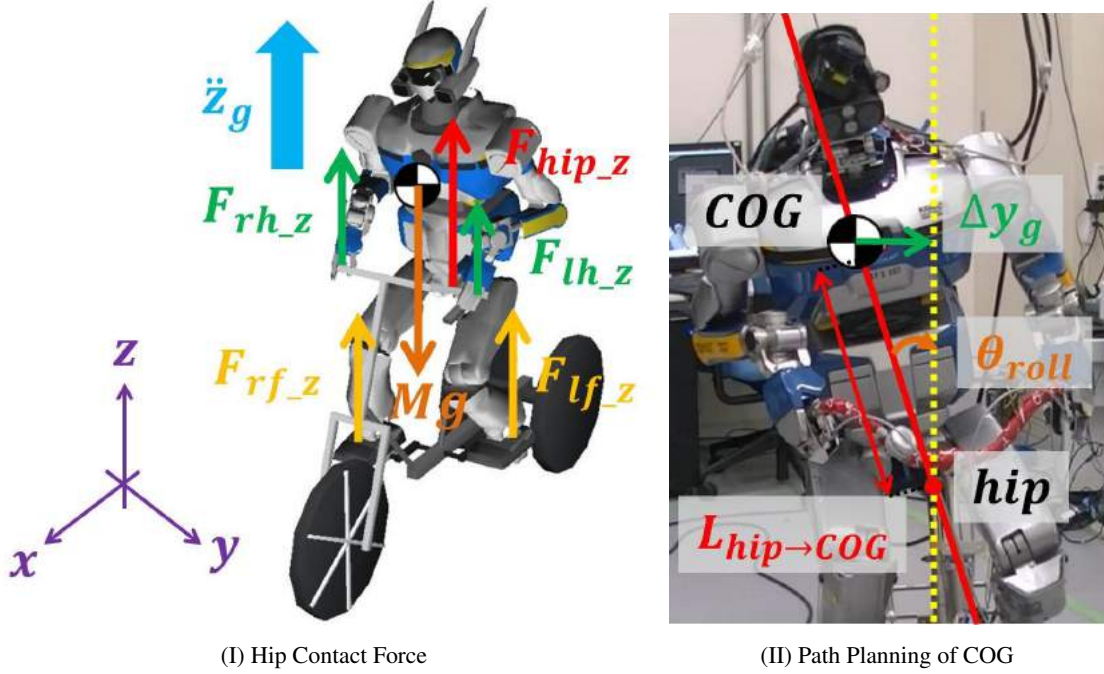


図 4.6: Strategies for Self Stability of Humanoid Robot [52]

(I) represents the relation between vertical hip contact force and the other contact forces, gravity, and inertia. (II) represents the recovery planning for COG.

できる。現在のヒューマノイドの体幹重心の y 成分位置を $y_g(n)$ とし、1 回のズレ補正動作完了後（割り込み処理型座り直し制御の完了後）の y 成分位置を $y_g(n+1)$ とするとき、ズレ補正量 Δy_g は以下で表現できる。

$$\Delta y_g = y_g(n+1) - y_g(n) \quad (4.4)$$

式 4.4 の Δy_g を式 4.3 に代入することで、重心の軌道が修正される更新則 $y_g(n) \rightarrow y_g(n+1)$ が次のように得られる。

$$y_g(n+1) := y_g(n) + L_{\text{hip} \rightarrow \text{COG}} \sin \theta_{\text{roll}}(n) \quad (4.5)$$

ここで、現在の Roll 方向の姿勢の傾斜角を $\theta_{\text{roll}}(n)$ と記述する。

次に、重心位置 $y_g(n)$ を軌道 $\rho(t; n)$ 、Roll 方向の姿勢の傾斜角 $\theta_{\text{roll}}(n)$ を探索変数 $\mathbf{x}(t; n)$ としたときのオンライン制御指標の定義式 2.5 における $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ 行列について考える。今回の例では式 4.5 の次元に関して、単一成分（Roll 方向）の次元から単一成分（ y 方向）の次元への 1 対 1 の 1 次元の写像であることから、 $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ 行列の次元は $\mathbb{R}^{1 \times 1}$

となる．式 2.7 のルールに基づいて $\rho(t; n)$ ($= y_g(n)$) に対する $\mathbf{x}(t; n)$ ($= \theta_{\text{roll}}(n)$) の単調性を調べる． $y_g(n)$ および $\theta_{\text{roll}}(n)$ の座標系設定を図 4.6 の向きを正とするとき， $y_g(n)$ の正の増加に伴い x まわりの傾斜角 $\theta_{\text{roll}}(n)$ は減少する方向に働くことから，式 2.7 の分類において単調減少に相当し $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ 行列には以下の 1 次元の値が割り当てられる．

$$\text{sgn}[\rho(t; n), \mathbf{x}(t; n)] = -1 \quad (4.6)$$

したがって，式 4.5 の更新式はオンライン制御指標の考えを用いた軌道修正の一般化更新則（式 2.14）の形式として以下で定式化することができる．

$$y_g(n+1) := y_g(n) - L_{\text{hip} \rightarrow \text{COG}} \text{sgn}[\rho(t; n), \mathbf{x}(t; n)] \sin(\theta_{\text{roll}}(n) - 0) \quad (4.7)$$

ここで探索目標 \mathbf{x}_d の値は，傾斜のない状態が理想的であることから $\mathbf{x}_d = 0$ に対応している．探索変数，探索目標，更新される軌道，一般化更新則の重み A のそれぞれの対応を表にまとめると以下となる．

表 4.3: 割り込み処理型座り直し制御におけるオンライン探索指標および軌道修正のパラメータの対応

探索変数	探索目標
$\theta_{\text{roll}}(n)$	0
更新される軌道	一般化更新則の重み A
$y_g(n)$	$L_{\text{hip} \rightarrow \text{COG}}$

ここで，重み行列 A は $L_{\text{hip} \rightarrow \text{COG}} > 0$ より正定値対称行列であることに留意する．

自己姿勢診断/補正行動システム

これまでに本項で述べた割り込み処理型座り直し制御の行動戦略におけるヒューマノイドの自己姿勢診断および補正行動システムの構成図を図 4.7 に示す．探索指標および軌道修正の観点からみた一連のアルゴリズムの処理の流れは以下の通りである．

1. ヒューマノイドの自己姿勢のズレ（オンライン探索指標 $f(t; n) = \theta_{\text{roll}}(n) - 0$ ）に対する診断系に基づく着座状態の搭乗姿勢のズレの検知（ヒューマノイド体内に搭載された IMU による体幹の Roll 傾斜の把握）

2. 双脚の両方に能動機能（オフライン探索指標による軌道計画）を適用したヒューマノイド臀部と搭乗型機器座面の間における接触の離脱

3. ヒューマノイド体幹の Roll 方向傾斜のエラー

（オンライン制御指標 $f_c(x(t;n)) = \text{sgn}[\rho(t;n), x(t;n)] \sin(\theta_{\text{roll}}(n) - 0)$

をフィードバックした搭乗姿勢の安定化

（重心位置の軌道修正 $y_g(n) \rightarrow y_g(n+1)$ ）

提案したシステムに基づく搭乗型機器の運転操作行動中における割り込み処理型座り直し制御の実機を用いた安定化の実現・検証に関して、後述の 7.4.1 項で示す。

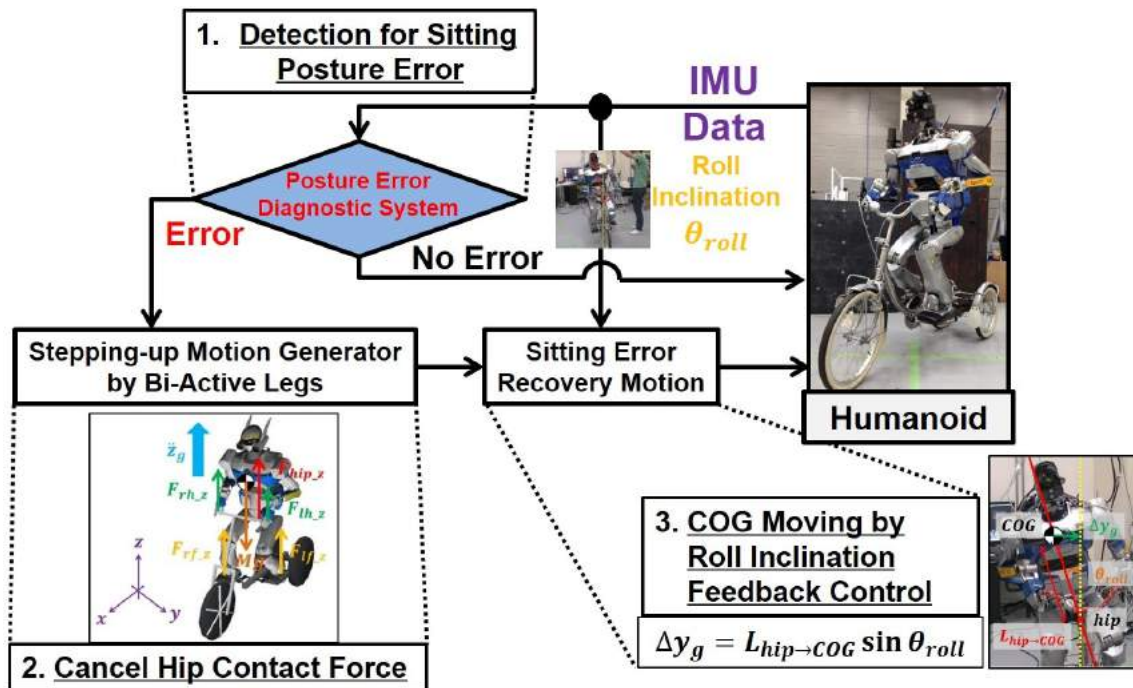


図 4.7: Posture Error Diagnostic and Self Recovery System Configuration for Humanoid Riding Tricycle

4.3.3 実時間並行処理型姿勢制御

前項で述べた割り込み処理型座り直し制御では、搭乗型機器上のヒューマノイドの搭乗姿勢のエラー（オンライン探索指標に相当する傾斜角）を体内搭載のIMUにより検知し、エラー情報をフィードバックしながら搭乗型機器の座面上に安定な姿勢で座り直す安定化行動制御法を提案した。割り込み処理型座り直し制御の実行処理は、搭乗型機器である三輪車のクランク角度が $\theta_c = k\pi$ (k : 整数) を満たすタイミングで実行される割り込み処理形式であり、運転操作の行動タスクを一時中断してから座り直し動作が生成される。そのため、円滑な運転操作行動において搭乗姿勢のズレが蓄積する度に、 $\theta_c = k\pi$ のタイミングで座り直しの割り込みが発生し、搭乗型機器の運転操作行動とヒューマノイドの姿勢安定化行動が互いに干渉し合う問題が存在する。特に、走行する路面環境の状態（傾斜勾配や凹凸など）が変化するとともに、ヒューマノイドが搭乗姿勢の傾きを検知し割り込み処理型座り直し制御を度々実行するようになり、速やかな運転操作が十分に達成されない状況等が想定される。

この問題に対して、新たに運転操作タスクと安定化行動タスクを同時に実行可能とする並行処理型の実時間バランス安定化制御を提案し、この制御法を「実時間並行処理型姿勢制御」と呼ぶ。

制御手法

実時間並行処理型姿勢制御は、二足歩行ロボットの歩行生成において同時に処理される安定化制御系の構成に着目し提案する。Kajita らが提案したヒューマノイドにおける線形倒立振子モデルのトラッキング制御法 [55][151][9] では、歩行パターン生成と共存する構成で安定化制御系が導入されている。本研究の目的であるヒューマノイドによる搭乗型機器の運転操作においては、二足歩行に代わり機器の操作に対して並行して処理されるヒューマノイドの搭乗姿勢の実時間安定化制御を図る。

実時間並行処理型姿勢制御が目的とする安定化対象はヒューマノイドの搭乗姿勢であり、文献 [55] で提案されている Chest posture control を活用した姿勢制御法が有効であると考え、この手法を搭乗型機器の運転操作に拡張する。Chest posture control [55] は、図 4.8 に示すような二足歩行型ヒューマノイドロボットの Chest フレーム（胸部リンク系） R_C に対する歩行パターンへの目標 R_C^d の追従を目的とした姿勢制御であり、センサデータをもとにした Chest フレームの Roll 成分と Pitch 成分の姿勢角を制御する。前項の割り込み処理型座り直し制御

で修正対象とした Roll 成分 (x 軸まわり) に関して, Chest フレームの Roll 成分の現在姿勢角を ϕ_C , 目標姿勢角を ϕ_C^d としたときの修正角 $\Delta\phi_C$ は文献 [55] より次式で定義される.

$$\Delta\phi_C = k_C (\phi_C^d - \phi_C) - \frac{1}{T_C} \Delta\phi_C \quad (4.8)$$

k_C は姿勢フィードバックのゲインであり, T_C は時定数を表す.

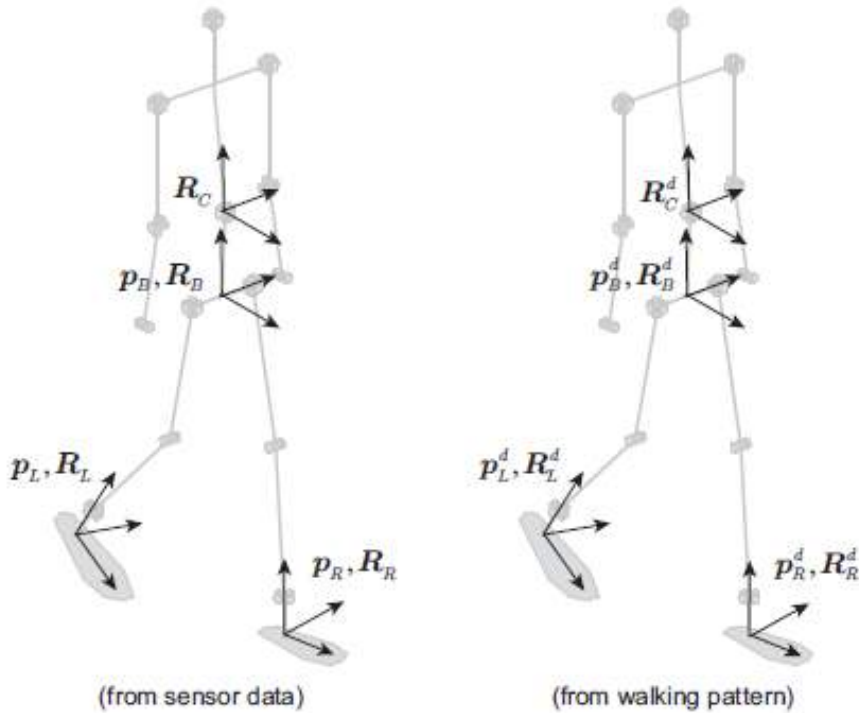


図 4.8: Frames for Stabilization Control [55]

次に, 式 4.8 から得られた Chest フレームの Roll 成分の修正角 $\Delta\phi_C$ の情報をもとに, どの点まわりで Chest フレームの目標値修正 $R_C^d \rightarrow R_C^{d*}$ (R_C^{d*} : 修正後の目標値) を行うかについて述べる. 二足歩行の場合における Chest posture control の姿勢角修正は, 重心まわりや二足歩行時の圧力中心となる支持脚を基準とした回転中心が設けられることが多い. 二足歩行の安定化と異なり, 搭乗型機器上のヒューマノイドは脚部に加えて腕部および臀部も圧力を受ける接触対象として拡張がなされる. このことから, 姿勢制御における回転中心は荷重が主要にかかる部位に設定することが望ましく, 例えば三輪車に搭乗するヒューマノイドに姿勢制御を適用する場合は以下の接触点の候補

1. ヒューマノイド右手-三輪車ハンドルグリップ

2. ヒューマノイド左手-三輪車ハンドルグリップ
3. ヒューマノイド右足-三輪車クランクペダル
4. ヒューマノイド左足-三輪車クランクペダル
5. ヒューマノイド臀部-三輪車サドル

のいずれかに設定する必要がある。本研究では4.3.1項の議論から、サドル座面（サドルノーズ）に対するヒューマノイド臀部の回転が搭乗姿勢のズレに影響し支配的であるとして、5の接触点を回転中心とした姿勢角修正を実現する。

以上の議論を整理した内容を図4.9に示す。図4.9における $\mathbf{R}_{rpy}^{hip}(\Delta\phi_C, 0, 0)$ は、5の接触点であるヒューマノイド臀部（Hip）を回転中心としてChestフレームを回転させる回転行列を表し、Roll方向に $\Delta\phi_C$ だけ回転させる変換行列である。これにより、Chestフレームの目標値修正 $\mathbf{R}_C^d \rightarrow \mathbf{R}_C^{d*}$ は以下の式で実行される。

$$\mathbf{R}_C^{d*} := \mathbf{R}_C^d \mathbf{R}_{rpy}^{hip}(\Delta\phi_C, 0, 0) \quad (4.9)$$

オンライン探索指標を利用した軌道修正によるアプローチ

制御手法で示した式（式4.8，式4.9）に基づき、オンライン探索指標の考えを取り入れた搭乗姿勢の安定化に対する軌道修正のアプローチを示す。はじめに式4.8の両辺にラプラス変換を施し、 $\Delta\phi_C$ のラプラス変換後の値 $\Delta\phi_C(s)$ について解いた後、逆変換より $\Delta\phi_C(t)$ を求めると以下が得られる。

$$\Delta\phi_C(t) = -\mathcal{L}^{-1} \left[\frac{k_C T_C}{1 + T_C s} (\phi_C(s) - \phi_C^d(s)) \right] \quad (4.10)$$

次に、目標値修正の式4.9は回転行列（変換行列）の積の形式で表現されており、離散時間系におけるChestフレームのRoll成分に関する目標姿勢角 $\phi_C^d(t)$ の更新式を簡略化すると以下で演算処理される。

$$\phi_C^d(t + \Delta t) := \phi_C^d(t) + \Delta\phi_C(t) \quad (4.11)$$

式4.10で得た $\Delta\phi_C(t)$ を式4.11に代入することで、新たに以下の目標値修正式が得られる。

$$\phi_C^d(t + \Delta t) := \phi_C^d(t) - \mathcal{L}^{-1} \left[\frac{k_C T_C}{1 + T_C s} (\phi_C(s) - \phi_C^d(s)) \right] \quad (4.12)$$

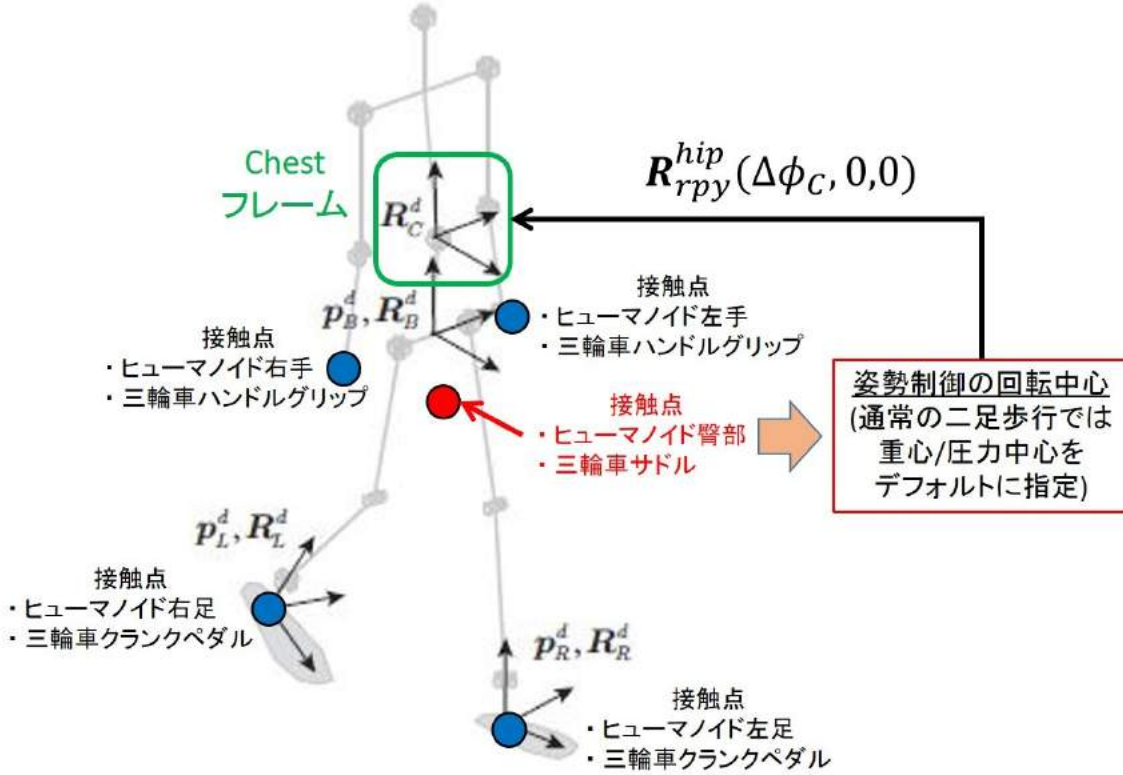


図 4.9: 搭乗型機器上のヒューマノイドに対する実時間並行処理型姿勢制御の回転中心の指定

さて、修正される軌道 $\rho(t; n)$ を搭乗姿勢における Chest フレームの目標 Roll 姿勢角 $\phi_C^d(t)$ 、探索変数 $\mathbf{x}(t; n)$ を Chest フレームの現在 Roll 姿勢角 $\phi_C(t)$ とすれば、 $\phi_C^d(t)$ と $\phi_C(t)$ は成分の方向が同一であり $\phi_C^d(t)$ が増加するとき $\phi_C(t)$ もまた増加を呈することから、式 2.7 の分類に基づいた $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ 行列は 1 次元で以下の値が振られる。

$$\text{sgn}[\rho(t; n), \mathbf{x}(t; n)] = 1 \quad (4.13)$$

したがって、目標値修正式 4.12 を $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ 行列を用いたオンライン制御指標による軌道修正の一般化更新則 (式 2.14) の形式で表すと以下で定式化される。

$$\phi_C^d(t + \Delta t) := \phi_C^d(t) - \mathcal{L}^{-1} \left[\frac{k_C T_C}{1 + T_C s} \text{sgn}[\rho(t; n), \mathbf{x}(t; n)] (\phi_C(s) - \phi_C^d(s)) \right] \quad (4.14)$$

式 4.14 より実時間並行処理型姿勢制御で対応する探索変数、探索目標、更新される軌道、一般化更新則の重み A は以下となる。

ここで、 $k_C > 0, T_C > 0$ であるため、 A は正定値対称行列の条件を満足している。

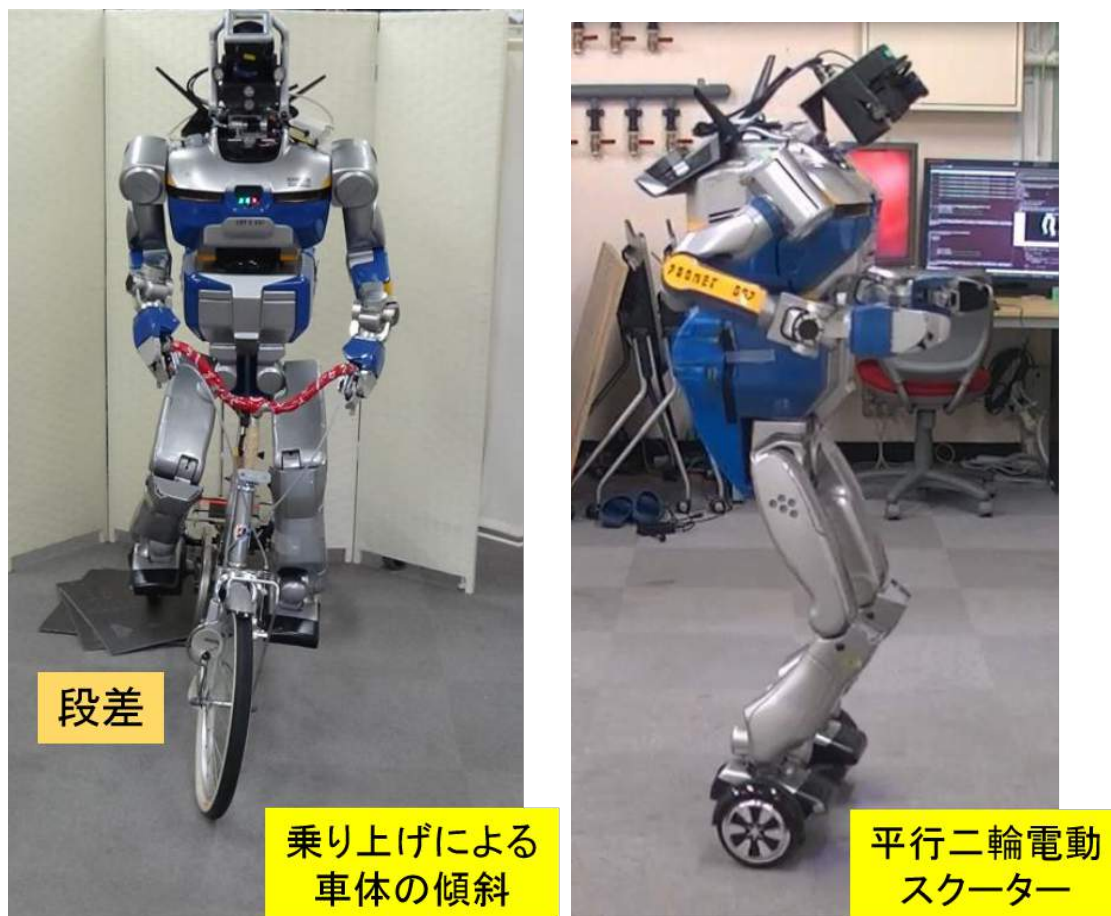
表 4.4: 実時間並行処理型姿勢制御におけるオンライン探索指標および軌道修正のパラメータの対応

探索変数	探索目標
$\phi_C(t)$	$\phi_C^d(t)$
更新される軌道	一般化更新則の重み A
$\phi_C^d(t)$	$\frac{k_C T_C}{1+T_C s}$

適用される挙動例

オンライン探索指標による搭乗姿勢の軌道修正に基づき実装された実時間並行処理型姿勢制御のロボット実機への適用検証例を示す。図 4.10 -(A) は三輪車の後輪が段差に乗り上げて三輪車全体が傾斜した状況下におけるヒューマノイドの搭乗姿勢の制御結果であり、搭乗したヒューマノイドが車体ごと傾斜した環境に応じて上体が水平となるよう実時間で修復を行いながらバランスを保持している。このように路面環境等により搭乗型機器自体が傾斜した場合、先に述べた割り込み処理型座り直し制御では、割り込み処理の適用範囲外 $\theta_c \neq k\pi$ の状態において不安定なまま運転操作行動を継続することとなり転倒する要因へと繋がる。これに対して、実時間並行処理型姿勢制御は運転操作行動と同時かつリアルタイムに割り込みなく実行されるため、図 4.10 -(A) に示した状況下の例などの路面状況に応じて迅速に対応することが可能である。実時間並行処理型姿勢制御による搭乗姿勢の安定化行動と搭乗型機器の運転操作行動を同時に実現した実証実験を後述の 7.3.1 項で示す。

上記以外の搭乗型機器上での搭乗姿勢の安定化の例として、図 4.10 -(B) にヒューマノイドによる平行二輪電動スクーター [58] の搭乗実験を示す。平行二輪電動スクーターは2つの車輪を有し、足部を乗せるボードの傾き度合いに応じて車輪を電動で駆動させる。搭乗するヒューマノイドの傾斜が大きければ大きいほど、平行二輪電動スクーターの移動量もその分大きくなり不安定となる。本実験では、姿勢制御の基準となる回転中心を両足の midpoint すなわち平行二輪電動スクーター車体の中心に設定した実時間並行処理型姿勢制御により、スクーター上のヒューマノイドの搭乗姿勢の傾斜を抑えて過度な移動を防いでいる。実時間でバランス安定化を並行処理として実現しているため、バランスを保持しながらの足部操作によるスクーターの運転操作に加えて腕部によるマニピュレーションに関しても同時実行が可能である。



(A) 三輪車全体が傾斜した状況下での姿勢制御

(B) 平行二輪電動スクーター [58] 上での姿勢制御

図 4.10: 異なる搭乗型機器への実時間並行処理型姿勢制御の適用例

4.3.4 割り込み処理型座り直し制御と実時間並行処理型姿勢制御の違い

これまでに提案した搭乗姿勢に対する2種類の安定化制御法（軌道修正法）を比較して長所と短所について整理する。

実時間並行処理型姿勢制御はリアルタイムかつマルチスレッドに搭乗姿勢の安定化行動が可能である一方で、割り込み処理型座り直し制御と比較して、搭乗型機器に着座したような多点接触環境では接触状態の解消を図る離脱遷移を伴わないため、接触した環境固定部位に対しての姿勢修正が困難な側面が存在する。例えば、三輪車に搭乗したヒューマノイドに実時間並行処理型姿勢制御を適用する場合は、Chest フレームを修正の対象としているが、サ

ドルと接触状態にある臀部が属する Base フレーム（骨盤リンク系）は接触拘束をもつため姿勢を自由に修正することは難しい。

他方で、割り込み処理型座り直し制御は行動のスレッドに割り込むような断続的な側面をもつが、座り直し動作を実行することで接触拘束を一時的に離脱させた Base フレーム（骨盤リンク系）も含めた姿勢修正を実現することができる。

これらの各制御方策がもつメリットとデメリットの側面を以下の表に整理する。

表 4.5: 割り込み処理型座り直し制御と実時間並行処理型姿勢制御におけるメリットとデメリット

	メリット	デメリット
割り込み処理型 座り直し制御	着座等の接触状態から一時的に離脱することで、Base フレーム（骨盤リンク系）も含めた搭乗姿勢の安定化が可能。	運転操作などの途中で割り込みが発生するため断続的な行動となり、また割り込み処理の範囲外では不安定性は解消されない。
実時間並行処理型 姿勢制御	搭乗姿勢の安定化行動と、運転操作などのその他の行動（マニピュレーション等）の共存が可能。	座面に着座した状態などの多点接触環境では接触の離脱は達成されないため、接触部位に対する姿勢修正は困難。

4.4 外界環境の認識に対する軌道修正法

本節では、ヒューマノイドと搭乗型機器を併合した系外に存在する外界の環境外乱に対するヒューマノイドと搭乗型機器の軌道修正法について述べる。

4.4.1 外界環境における外乱要因と認識行動方策

搭乗型機器および搭乗したヒューマノイドからなる系以外の環境を外界環境と定義する。外界環境には多種多様な系が存在するが、そのなかでも特に搭乗型機器の運転操作時において妨げとなる障害物や地形等を外乱要因とする。また、外乱要因から構成される環境系を障害環境とよぶものとする。

本研究では、障害環境に存在する外乱要因を

- 視野が可到達な範囲の大型障害物
- 視野が不可到達な範囲の小型障害物

に切り分けて、その行動方策を計画する．大型障害物および小型障害物の具体的な例については表 4.1，図 4.2 において既に示した．

外界に存在する障害環境に対する認識行動方策として，大型・小型それぞれの障害物に関する認識法を分別する．ヒューマノイドの視野が可到達かつ視覚認識可能な範囲内における大型障害物の認識は，ヒューマノイドの頭部に搭載されたステレオカメラを用いたビジョン認識により把握する．一方で，視野が不可到達または視覚認識困難な範囲に存在する小型障害物は，搭乗型機器を介してヒューマノイドが検出する力学的負荷に基づいた認識法により，障害物の存在を把握する戦略をとる．

これらの各障害環境に対する認識アプローチを図 4.11 にまとめた．

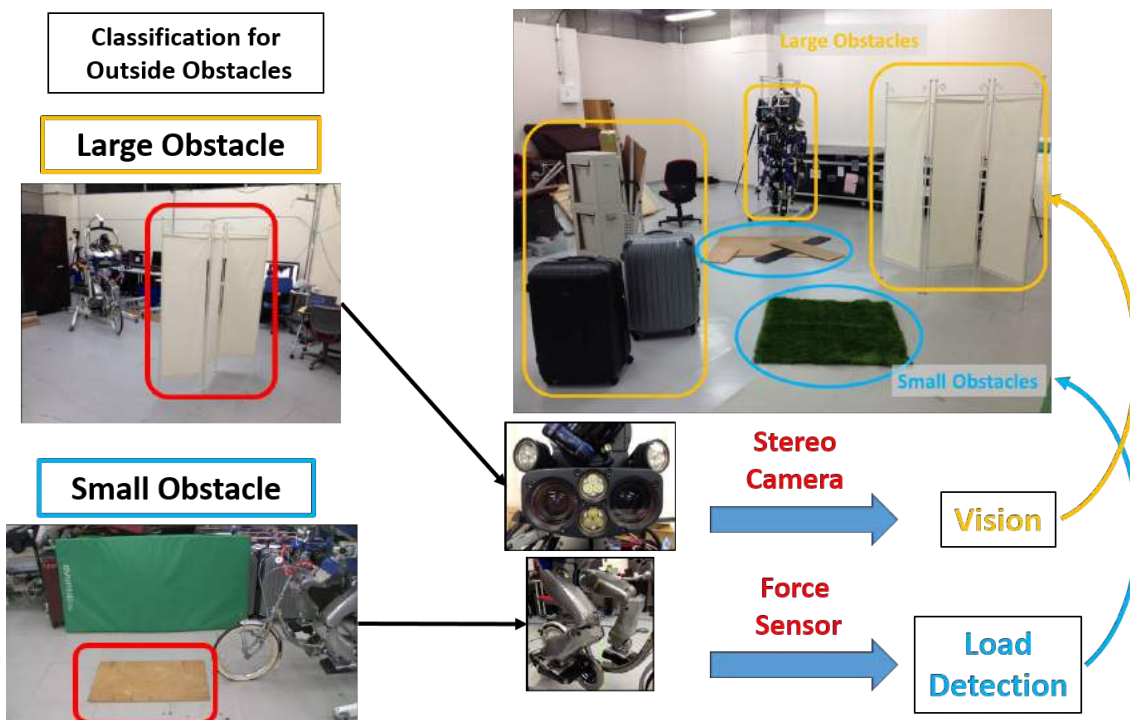


図 4.11: Classification for Outside Obstacles and Perception Approach

4.4.2 大型障害環境に対する認識行動

大型障害物の認識法

大型障害物はロボットの視野が可到達な範囲において存在する障害物のため、ロボットに搭載されたカメラデバイスを利用した視覚認識が有効である。ロボットによる障害物認識の研究 [152][153][154] は幅広くおこなわれており、ヒューマノイドにおいては頭部に搭載されたステレオカメラによる認識や、LRF (Laser Range Finder) を用いた距離・測域情報の認識、これらを組み合わせた [155] のようなカメラデバイスを活用した認識方法などが一般的である。

本研究では、大型障害物の認識法として、ステレオカメラから計測された三次元の点群データ (Point Cloud) の情報より構成される障害物体を、ロボットがもっている白地図上に配置したローカルコストマップの形成手法 [156][157] を採用する。認識・作成したローカルコストマップの情報に基づきヒューマノイドによる搭乗型機器の操作行動生成を計画する。

大型障害環境における行動生成法

作成したローカルコストマップが、ヒューマノイドがオンラインで所有している大型障害環境のマップにあたる。大型障害物の存在は点群データとして三次元空間上で表されるが、ローカルコストマップはそれら障害物からの仮想ポテンシャル場をもった二次元平面で構成される。そのため、地図平面におけるゴールとなる目標位置・姿勢を操作インターフェースより明示的に与えたとき、ヒューマノイドが搭乗型機器を操作してその地点に到達するための行動戦略が要求される。行動戦略を目標までの経路計画と、経路に追従するための操縦戦略に分別して考える。

経路計画 経路計画では、ロボットがもっているローカルコストマップを入力、ロボットの2次元の自己位置における目標速度・角速度を出力とする行動戦略をとる。平面内におけるローカルナビゲーションシステムとして、Dynamic Window Approach (DWA) [158] を利用したローカルプランナを構成する。ローカルプランナはROS (Robot Operating System) [140][141] より提供される move_base ライブラリ [79] を利用する。move_base におけるシステム構成図は、図 4.12 のように表され、本提案手法においては図の赤囲で示されたノードお

よびセンサデータを順次適用していく。これにより、ロボットの目標速度・角速度(図 4.12 における”cmd_vel”)を得る。

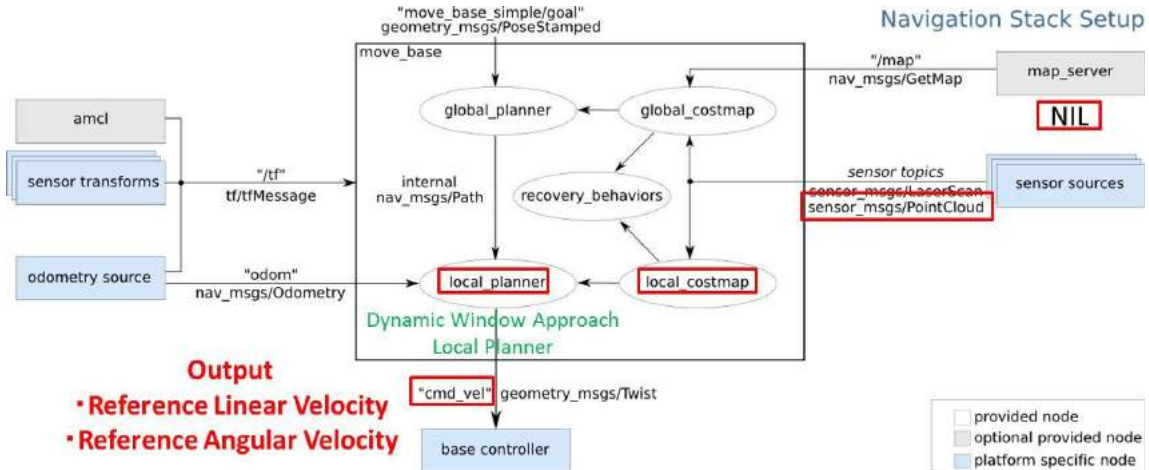


図 4.12: Global and Local Planner Package “move_base” [79] Provided by ROS (Robot Operating System) [140]

三輪車に対する操縦戦略 経路計画に基づいたヒューマノイドによる搭乗型機器への操縦戦略として、三輪車への目標操作量(目標ハンドル角度・目標クランク角度)の軌道修正を考える。三輪車の操縦行動では、経路計画で得られた平面地図上での目標速度・角速度を入力、三輪車の目標ハンドル操舵角・目標クランク操作角を出力とした行動戦略をとる。図 4.13 のような平面内における三輪車のオドメトリモデルを考える。三輪車は 1 後輪駆動、1 ステアリングとする。三輪車の一般化速度は $(\dot{x} \ \dot{y} \ \dot{\theta} \ \dot{\phi})^T$ で表されており、それぞれ \dot{x} , \dot{y} , $\dot{\theta}$: ワールド座標系における三輪車の x , y , θ 成分の速度, $\dot{\phi}$: 三輪車のハンドル角速度を表す。また、三輪車の制御入力を $(v_b \ \omega_h)^T$ で表し, v_b : 後輪の並進駆動速度入力, ω_h : ハンドル角速度入力とする。

このとき、三輪車の一般化速度と制御入力の関係は式 4.15 となる。

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \end{pmatrix} = \begin{pmatrix} \cos \theta & 0 \\ \sin \theta & 0 \\ \frac{\tan \phi}{L} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v_b \\ \omega_h \end{pmatrix} \quad (4.15)$$

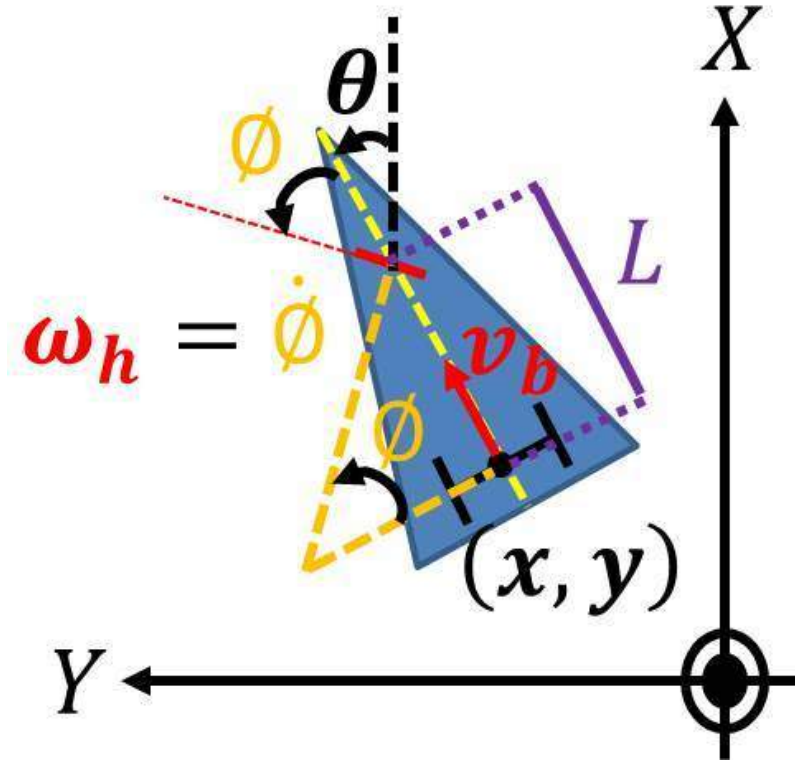


図 4.13: Tricycle Odometry on 2D Plane [53]
 State of velocity is $(\dot{x} \ \dot{y} \ \dot{\theta} \ \dot{\phi})^T$, and control input is $(v_b \ \omega_h)^T$.

L はホイールベース長である。また、制御入力 $(v_b \ \omega_h)^T$ を三輪車のクランク角速度，操舵ハンドル角速度 $(\omega_c \ \omega_h)^T$ に変換する式を与えると

$$\begin{aligned} \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \end{pmatrix} &= \begin{pmatrix} \cos \theta & 0 \\ \sin \theta & 0 \\ \frac{\tan \phi}{L} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{D_c}{2\pi} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \omega_c \\ \omega_h \end{pmatrix} \\ &= \begin{pmatrix} \frac{D_c \cos \theta}{2\pi} & 0 \\ \frac{D_c \sin \theta}{2\pi} & 0 \\ \frac{D_c \tan \phi}{2\pi L} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \omega_c \\ \omega_h \end{pmatrix} \end{aligned} \quad (4.16)$$

となる。 D_c はクランク1周期に対して三輪車が滑らずに進行する距離であり，三輪車固有のパラメータである。

さて，ローカルプランナから求められる目標速度・角速度 $\dot{x}^{\text{ref}}, \dot{y}^{\text{ref}}, \dot{\theta}^{\text{ref}}$ についても，式4.16は

適用可能であり、またローカルプランナでは座標系が三輪車固定のローカル座標系となるため、 $\theta = 0$ を常に満たす。このとき、式 4.16 は次式で書き表すことができる。

$$\begin{pmatrix} \dot{x}^{\text{ref}} \\ \dot{y}^{\text{ref}} \\ \dot{\theta}^{\text{ref}} \\ \dot{\phi}^{\text{ref}} \end{pmatrix} = \begin{pmatrix} \frac{D_c}{2\pi} & 0 \\ 0 & 0 \\ \frac{D_c \tan \phi^{\text{ref}}}{2\pi L} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \omega_c^{\text{ref}} \\ \omega_h^{\text{ref}} \end{pmatrix} \quad (4.17)$$

式 4.17 の第 1 行、第 3 行、第 4 行より三輪車の目標クランク角速度 ω_c^{ref} 、目標ハンドル操舵角速度 ω_h^{ref} 、ならびに目標ハンドル操舵角 ϕ^{ref} は次のように求めることができる (第 2 行は $\dot{y}^{\text{ref}} = 0$ を満たしている)。

$$\omega_c^{\text{ref}} = \frac{2\pi \dot{x}^{\text{ref}}}{D_c} \quad (4.18)$$

$$\omega_h^{\text{ref}} = \dot{\phi}^{\text{ref}} \quad (4.19)$$

$$\begin{aligned} \phi^{\text{ref}} &= \arctan\left(\frac{2\pi L \dot{\theta}^{\text{ref}}}{D_c \omega_c^{\text{ref}}}\right) \\ &= \arctan\left(\frac{L \dot{\theta}^{\text{ref}}}{\dot{x}^{\text{ref}}}\right) \end{aligned} \quad (4.20)$$

以上より、出力となる目標ハンドル角、目標クランク角 $(\theta_h^{\text{ref}} \theta_c^{\text{ref}})^{\text{T}}$ は、プランナから得られる入力 $(\dot{x}^{\text{ref}} \dot{\theta}^{\text{ref}})^{\text{T}}$ を用いて以下で決定される。

$$\begin{aligned} \begin{pmatrix} \theta_h^{\text{ref}} \\ \theta_c^{\text{ref}} \end{pmatrix} &= \begin{pmatrix} \phi^{\text{ref}} \\ \theta_c + \omega_c^{\text{ref}} dt \end{pmatrix} \\ &= \begin{pmatrix} \arctan\left(\frac{L \dot{\theta}^{\text{ref}}}{\dot{x}^{\text{ref}}}\right) \\ \theta_c + \frac{2\pi \dot{x}^{\text{ref}}}{D_c} dt \end{pmatrix} = \mathbf{F}(\dot{x}^{\text{ref}}, \dot{\theta}^{\text{ref}}) \end{aligned} \quad (4.21)$$

θ_c は現在時刻 t でのクランク角度であり、 dt は離散時間間隔を表す。式 4.21 の目標値に従って、第 3 章で述べた運転操作に対する軌道修正を実行することで大型障害環境下での行動生成が達成される。

ナビゲーション情報を利用したオンライン探索指標に基づく軌道修正

大型障害環境の認識行動における軌道修正法の立ち位置 大型障害環境においては、2次元平面上での目標値 $(\dot{x}^{\text{ref}} \dot{\theta}^{\text{ref}})^{\text{T}}$ を搭乗型機器への目標操作量 $(\theta_h^{\text{ref}} \theta_c^{\text{ref}})^{\text{T}}$ に変換するアプローチを提案している。このときの $(\dot{x}^{\text{ref}} \dot{\theta}^{\text{ref}})^{\text{T}} \rightarrow (\theta_h^{\text{ref}} \theta_c^{\text{ref}})^{\text{T}}$ の変換は写像として考えることができ、2.5 節で論じたオンライン探索指標に基づく軌道修正のアプローチとして捉えることができ

る。軌道修正による写像変換前の状態 $(\dot{x}^{\text{ref}} \ \dot{\theta}^{\text{ref}})^{\text{T}}$ は探索変数の要素としてみることができ、これらのパラメータはプランナが教示する一種のナビゲーション情報と言える。写像変換後の状態 $(\theta_h^{\text{ref}} \ \theta_c^{\text{ref}})^{\text{T}}$ はリファレンスとなる軌道であり、プランナ教示によるナビゲーション情報を基に実時間で逐次更新・修正されていることが式 4.21 より理解できる。

したがって、軌道修正の一般化更新則における探索変数はナビゲーション情報 $(\dot{x}^{\text{ref}} \ \dot{\theta}^{\text{ref}})^{\text{T}}$ であり、探索目標はゴールにおいて $(0 \ 0)^{\text{T}}$ となる。また、修正される軌道は搭乗型機器への目標操作量 $(\theta_h^{\text{ref}} \ \theta_c^{\text{ref}})^{\text{T}}$ が相当する。

これまでに述べた大型障害環境に対する一連の認識行動戦略の処理を表した図が図 4.14 であり、この処理系においてナビゲーション情報を利用したオンライン探索指標に基づく軌道修正の処理ブロックは図 4.14 内の赤枠で示した囲み部分に位置する。図 4.14 の処理系で示されている通り、コストマップの作成やローカルプランナなどによる処理を経て得られたナビゲーション情報をもとにした搭乗型機器への目標操作量の軌道修正が行われている。軌道修正により決定した目標ハンドル角度や目標クランク角度に追従するヒューマノイドのマニピュレーション戦略については、第3章で述べた能動機能と受動機能のマニピュレータ軌道制御が担う。

軌道修正の更新則の定式化 図 4.14 で示した軌道修正処理の立ち位置を踏まえて、軌道修正の更新則（式 2.14）に準拠した定式化・各パラメータの対応を述べる。はじめに、軌道修正における変換写像に対応する式 4.21 を離散時間系の更新式として記述し直すと以下で表される。

$$\begin{pmatrix} \theta_h^{\text{ref}}(t + \Delta t) \\ \theta_c^{\text{ref}}(t + \Delta t) \end{pmatrix} := \begin{pmatrix} \arctan\left(\frac{L\dot{\theta}^{\text{ref}}(t)}{\dot{x}^{\text{ref}}(t)}\right) \\ \theta_c(t) + \frac{2\pi\Delta t}{D_c}\dot{x}^{\text{ref}}(t) \end{pmatrix} \quad (4.22)$$

t は現在時刻であり、 dt をサンプリング時間 Δt に表記し直すと次ステップでの時刻は $t + \Delta t$ と書ける。

次に、式 4.22 から探索変数となるナビゲーション情報 $(\dot{\theta}^{\text{ref}}(t) \ \dot{x}^{\text{ref}}(t))^{\text{T}}$ のベクトルを括り出す。式 4.22 の1行目は $\dot{\theta}^{\text{ref}}(t)$, $\dot{x}^{\text{ref}}(t)$ に関して非線形であるため、ここではテイラー展開を用いた1次近似線形化 [159] を図る。 $(\dot{\theta}^{\text{ref}}(t) \ \dot{x}^{\text{ref}}(t))^{\text{T}}$ の平衡点を $(\dot{\theta}_*^{\text{ref}} \ \dot{x}_*^{\text{ref}})^{\text{T}}$ とするとき、式 4.22 の

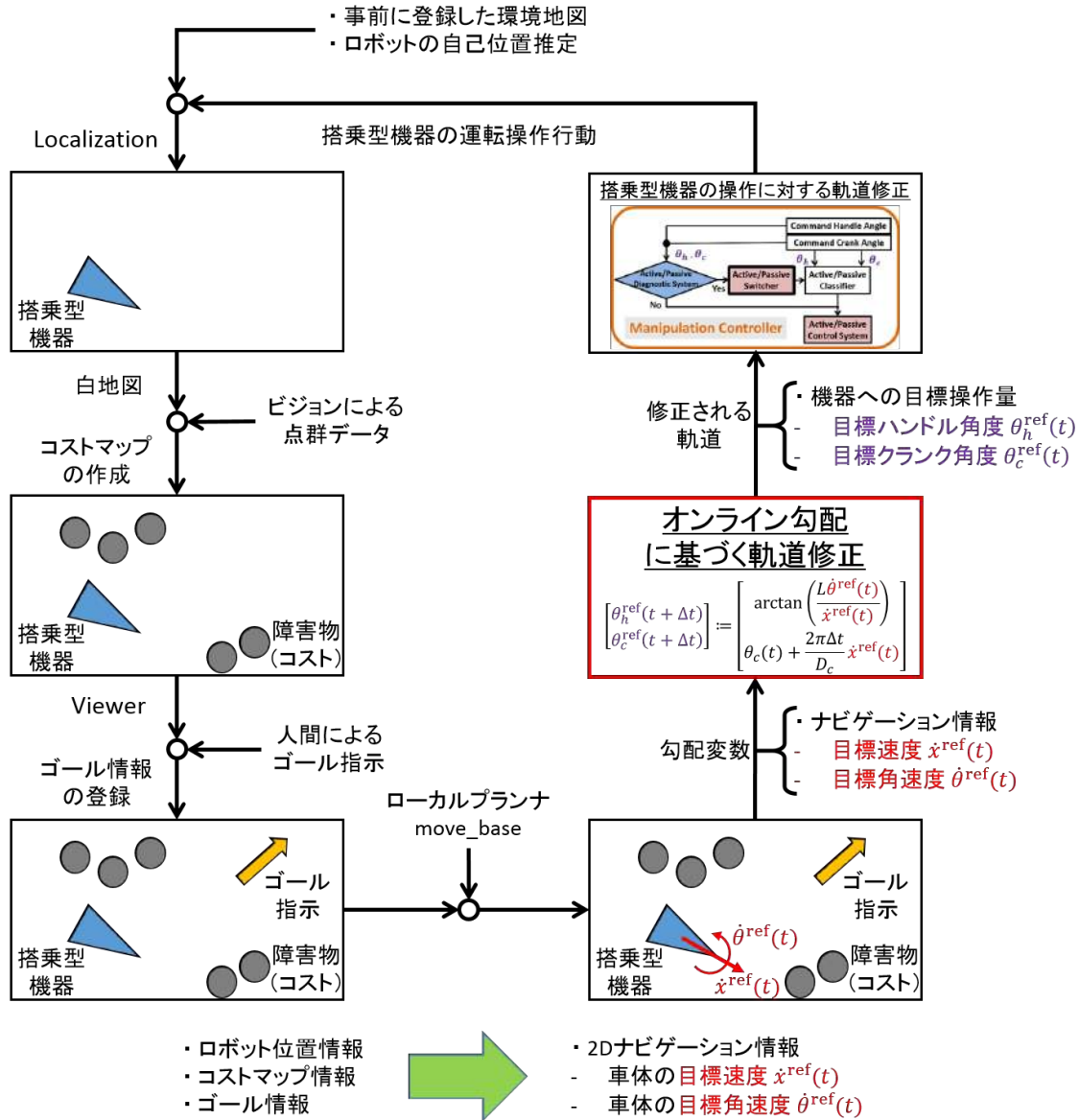


図 4.14: 大型障害環境に対する認識行動処理系
 ロボット視覚によるナビゲーション情報を利用した軌道修正処理の立ち位置は、赤枠で囲まれた処理ブロックに該当する。

1行目の $\theta_h^{\text{ref}}(t + \Delta t)$ は $(\dot{\theta}_*^{\text{ref}} \dot{x}_*^{\text{ref}})^T$ まわりで次式のように線形化される。

$$\begin{aligned} \theta_h^{\text{ref}}(t + \Delta t) &\simeq \theta_h^{\text{ref}}(t + \Delta t) \Big|_{\dot{\theta}_*^{\text{ref}}, \dot{x}_*^{\text{ref}}} \\ &+ \frac{\partial \theta_h^{\text{ref}}(t + \Delta t)}{\partial \dot{\theta}_*^{\text{ref}}(t)} \Big|_{\dot{\theta}_*^{\text{ref}}, \dot{x}_*^{\text{ref}}} (\dot{\theta}_*^{\text{ref}}(t) - \dot{\theta}_*^{\text{ref}}) + \frac{\partial \theta_h^{\text{ref}}(t + \Delta t)}{\partial \dot{x}_*^{\text{ref}}(t)} \Big|_{\dot{\theta}_*^{\text{ref}}, \dot{x}_*^{\text{ref}}} (\dot{x}_*^{\text{ref}}(t) - \dot{x}_*^{\text{ref}}) \\ &\simeq \frac{L \dot{x}_*^{\text{ref}}}{\dot{x}_*^{\text{ref}2} + L^2 \dot{\theta}_*^{\text{ref}2}} \dot{\theta}_*^{\text{ref}}(t) - \frac{L \dot{\theta}_*^{\text{ref}}}{\dot{x}_*^{\text{ref}2} + L^2 \dot{\theta}_*^{\text{ref}2}} \dot{x}_*^{\text{ref}}(t) + \theta_{h*} \end{aligned} \quad (4.23)$$

ここで、平衡点 $(\dot{\theta}_*^{\text{ref}} \dot{x}_*^{\text{ref}})^T$ に関して整理した定数項を θ_{h*} とする。

平衡点を原点 $(\dot{\theta}_*^{\text{ref}} \dot{x}_*^{\text{ref}})^T = (0 \ 0)^T$ としたとき、式 4.23 の項の分母は 0 となり定義されないため、原点近傍の点 $(0 \ \delta \dot{x}_*^{\text{ref}})^T$ ($\delta \dot{x}_*^{\text{ref}} > 0$ は十分に微小な正值) を平衡点とするとき、式 4.23 は以下となる。

$$\theta_h^{\text{ref}}(t + \Delta t) \simeq \frac{L}{\delta \dot{x}_*^{\text{ref}}} \dot{\theta}_*^{\text{ref}}(t) \quad (4.24)$$

このとき式 4.23 の定数項 θ_{h*} は、平衡点 $(0 \ \delta \dot{x}_*^{\text{ref}})^T$ において 0 となる。

以上より、式 4.22 と式 4.24 から $(\dot{\theta}_*^{\text{ref}}(t) \dot{x}_*^{\text{ref}}(t))^T$ に関して以下で近似線形化される。

$$\begin{pmatrix} \theta_h^{\text{ref}}(t + \Delta t) \\ \theta_c^{\text{ref}}(t + \Delta t) \end{pmatrix} := \begin{pmatrix} 0 \\ \theta_c(t) \end{pmatrix} + \begin{pmatrix} \frac{L}{\delta \dot{x}_*^{\text{ref}}} & 0 \\ 0 & \frac{2\pi \Delta t}{D_c} \end{pmatrix} \begin{pmatrix} \dot{\theta}_*^{\text{ref}}(t) \\ \dot{x}_*^{\text{ref}}(t) \end{pmatrix} \quad (4.25)$$

最後に、 $\text{sgn}[\boldsymbol{\rho}(t; n), \mathbf{x}(t; n)]$ 行列 (式 2.6) について考える。軌道を $\boldsymbol{\rho}(t; n) = (\theta_h^{\text{ref}}(t) \theta_c^{\text{ref}}(t))^T$ 、探索変数を $\mathbf{x}(t; n) = (\dot{\theta}_*^{\text{ref}}(t) \dot{x}_*^{\text{ref}}(t))^T$ とするとき、 $\text{sgn}[\boldsymbol{\rho}(t; n), \mathbf{x}(t; n)]$ 行列は $\mathbb{R}^{2 \times 2}$ 次元である。

$\text{sgn}[\boldsymbol{\rho}(t; n), \mathbf{x}(t; n)]$ 行列の 1 行 1 列目の要素に相当する $\theta_h^{\text{ref}}(t)$ に対しての $\dot{\theta}_*^{\text{ref}}(t)$ の単調性を調べる。 $\theta_h^{\text{ref}}(t)$ は三輪車の目標ハンドル角度、 $\dot{\theta}_*^{\text{ref}}(t)$ は 2 次元平面上での三輪車車体の Yaw まわりの目標角速度を表すが、単調性を調べるにあたり以下の 2 つの変数を導入する。

$\dot{\theta}_{nav}^{\text{ref}}$ ：ナビゲーションが教示する (車体の) 角速度

$\dot{\theta}(t)$ ：実際に出力される (車体の) 角速度

$\dot{\theta}_{nav}^{\text{ref}}$ は図 4.12 のローカルプランナから与えられる指針となるナビゲーション情報であり、 $\dot{\theta}_*^{\text{ref}}(t)$ の単調性を調べるためにここでは固定した一定値とする。 $\dot{\theta}(t)$ が実際の三輪車車体に出力される Yaw まわりの角速度である。

このとき、 $\dot{\theta}_*^{\text{ref}}(t)$ は実時間で決定される変数であり、 $\dot{\theta}_{nav}^{\text{ref}}$ および $\dot{\theta}(t)$ との関係性は以下で与えられる。

$$\dot{\theta}_*^{\text{ref}}(t) = \dot{\theta}_{nav}^{\text{ref}} - \dot{\theta}(t) \quad (4.26)$$

式 4.26 の関係性をグラフとして図示したものが図 4.15 である。グラフの横軸に相当する目標ハンドル角度 $\theta_h^{\text{ref}}(t)$ の値が増加するとき、車体の旋回半径が小さくなり実際に出力される車体の角速度 $\dot{\theta}(t)$ もまた増加傾向を示す。 $\dot{\theta}_{\text{nav}}^{\text{ref}}$ は先ほど述べた通り一定値であるため、式 4.26 の関係性を用いると $\dot{\theta}^{\text{ref}}(t)$ はグラフの赤線で示されるように単調減少を呈する。したがって、式 2.7 の分類から $\text{sgn}[\theta_h^{\text{ref}}(t), \dot{\theta}^{\text{ref}}(t)] = -1$ が割り当てられる。

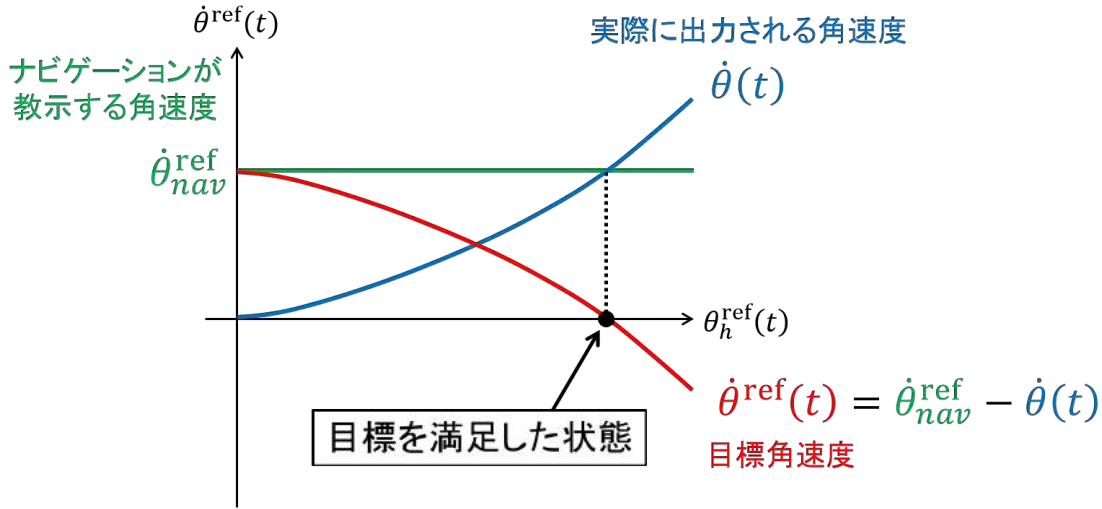


図 4.15: 目標ハンドル角度 $\theta_h^{\text{ref}}(t)$ に対する車体の目標角速度 $\dot{\theta}^{\text{ref}}(t)$ の単調性
 $\dot{\theta}_{\text{nav}}^{\text{ref}}$ が一定のとき、実際に出力される車体の角速度 $\dot{\theta}(t)$ は単調増加を表し、一方で車体の目標角速度 $\dot{\theta}^{\text{ref}}(t)$ は単調減少を呈する。

$\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ 行列の 1 行 2 列目の $\theta_h^{\text{ref}}(t)$ に対する $\dot{x}^{\text{ref}}(t)$ の挙動は、ハンドルをどれだけ操舵させても車体の並進速度には非依存となるため $\text{sgn}[\theta_h^{\text{ref}}(t), \dot{x}^{\text{ref}}(t)] = 0$ である。 $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ 行列の 2 行 1 列目にあたる $\theta_c^{\text{ref}}(t)$ に対する $\dot{\theta}^{\text{ref}}(t)$ の挙動は、ハンドルの目標角度 $\theta_h^{\text{ref}}(t)$ の符号に応じて図 4.16 に示すように $\dot{\theta}^{\text{ref}}(t)$ が振動的になり不安定な単調性を呈するため $\text{sgn}[\theta_c^{\text{ref}}(t), \dot{\theta}^{\text{ref}}(t)] = 0$ となる。 $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ 行列の 2 行 2 列目は $\theta_c^{\text{ref}}(t)$ に対する $\dot{x}^{\text{ref}}(t)$ の挙動であり、クランクの目標操作角が大きければ大きいほど、実際に出力される車体の並進速度 $\dot{x}(t)$ は増加し、図 4.15 で観察した挙動と同様に $\text{sgn}[\theta_c^{\text{ref}}(t), \dot{x}^{\text{ref}}(t)] = -1$ が割り当てられる。

以上より、 $\text{sgn}[\rho(t; n), \mathbf{x}(t; n)]$ 行列は次式で与えられる。

$$\text{sgn}[\rho(t; n), \mathbf{x}(t; n)] = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -E \quad (4.27)$$

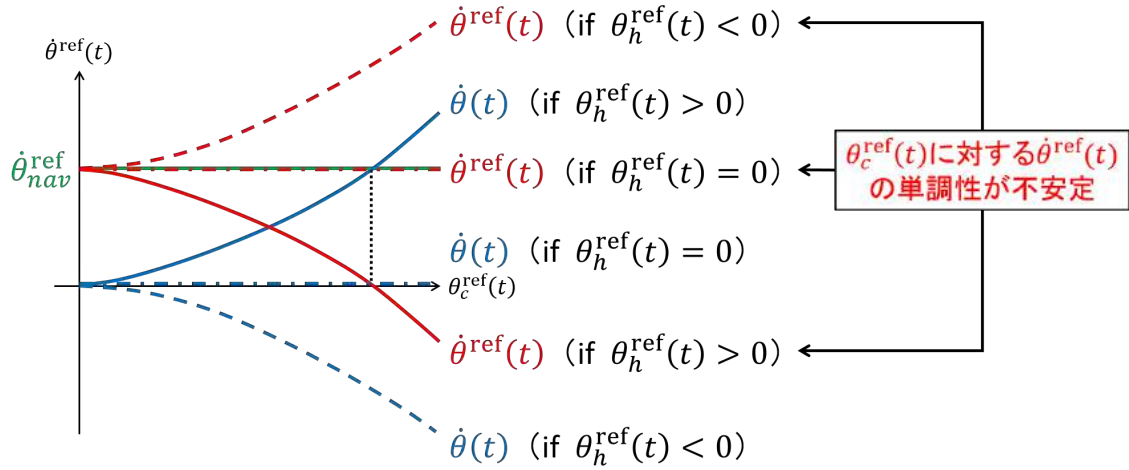


図 4.16: 目標クランク角度 $\theta_c^{\text{ref}}(t)$ に対する車体の目標角速度 $\dot{\theta}^{\text{ref}}(t)$ の単調性
 目標クランク角度 $\theta_c^{\text{ref}}(t)$ の増加に対して車体の目標角速度 $\dot{\theta}^{\text{ref}}(t)$ は、目標ハンドル角度 $\theta_h^{\text{ref}}(t)$ の状態に応じて単調性が異なるため不安定となる。

式 4.25，式 4.27 から，オンライン探索指標（オンライン制御指標）に基づく軌道修正の更新則は以下で定式化される。

$$\begin{pmatrix} \theta_h^{\text{ref}}(t + \Delta t) \\ \theta_c^{\text{ref}}(t + \Delta t) \end{pmatrix} := \begin{pmatrix} 0 \\ \theta_c(t) \end{pmatrix} - \begin{pmatrix} \frac{L}{\delta \dot{x}_*^{\text{ref}}} & 0 \\ 0 & \frac{2\pi\Delta t}{D_c} \end{pmatrix} \mathbf{sgn}[\rho(t; n), \mathbf{x}(t; n)] \left\{ \begin{pmatrix} \dot{\theta}^{\text{ref}}(t) \\ \dot{x}^{\text{ref}}(t) \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \quad (4.28)$$

一般化更新則（式 2.14）とのパラメータの対応は以下の通りである。

表 4.6: 視覚ナビゲーション情報に基づくオンライン探索指標および軌道修正のパラメータの対応

探索変数	探索目標
$(\dot{\theta}^{\text{ref}}(t) \dot{x}^{\text{ref}}(t))^{\text{T}}$	$(0 \ 0)^{\text{T}}$
更新される軌道	一般化更新則の重み A
$(\theta_h^{\text{ref}}(t) \theta_c^{\text{ref}}(t))^{\text{T}}$	$\begin{pmatrix} \frac{L}{\delta \dot{x}_*^{\text{ref}}} & 0 \\ 0 & \frac{2\pi\Delta t}{D_c} \end{pmatrix}$

探索目標 $(0 \ 0)^{\text{T}}$ においてはゴールの状態を表している。また， $\frac{L}{\delta \dot{x}_*^{\text{ref}}} > 0, \frac{2\pi\Delta t}{D_c} > 0$ より A は正定値対称行列である。式 4.28 は，一般化更新則で定義した式 2.14 と比べて軌道修正前の項が $(\theta_h^{\text{ref}}(t) \theta_c^{\text{ref}}(t))^{\text{T}}$ とはならず $(0 \ \theta_c(t))^{\text{T}}$ で定式化されている。この定式化の結果より，次の

特徴が定量的に示された。

- 目標ハンドル角度の修正においては、前回（前時刻）の軌道は用いずに 0[deg] の基準点からの修正を行っている。
- 目標クラック角度の修正においては、目標値ではなく現在値のクラック角度から修正を行うことが可能である。

大型障害環境に対する認識行動システム

これまでに論じた大型障害環境における認識法・行動生成計画法・軌道修正戦略などを基に、図 4.17 に示すような大型障害環境認識行動システムを構成する。

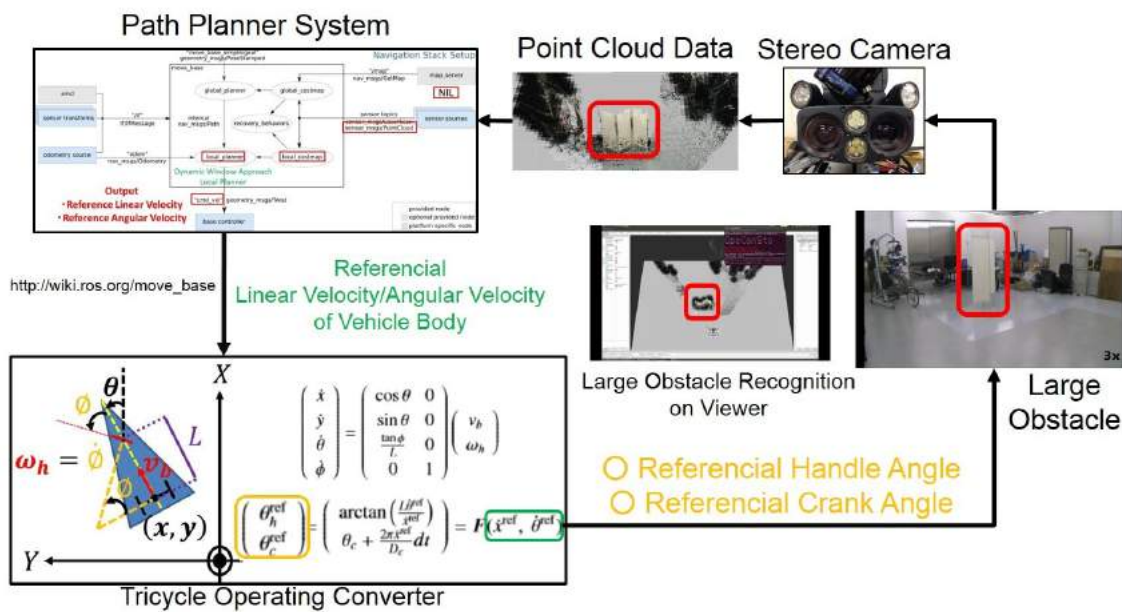


図 4.17: Recognition and Planning System for Large Obstacles

図 4.17 の大型障害環境認識行動システムのアルゴリズムを以下に示す。

1. ヒューマノイド頭部に搭載されたステレオカメラによる大型障害物の点群データの取得。
2. 構成された障害環境の点群データよりコストマップを形成。

3. ローカルプランナ `move_base` [79] に基づいた搭乗型機器車体の目標速度・目標角速度を算出.
4. 提案した軌道修正法に基づく車体の目標速度・目標角速度から搭乗型機器の目標操作量への変換.
5. 第3章で提案した能動機能と受動機能のマニピュレータ軌道制御による目標操作量に対する実機実現.
6. 手順1へ帰還.

先に示した図 4.14 や構成した図 4.17 の認識行動システムでは、閉ループの帰還処理系として実装されており、メインとなる運転操作の行動生成に対して別の認識行動生成が割り込むことはなく、割り込みによる運転操作の逐次的な中断等は発生しない。したがって、運転操作と並行して外界環境の状況に応じた行動生成が実現可能である。

大型障害環境における運転操作行動実験

図 4.17 で構築した認識行動システムを利用した大型障害環境下でのヒューマノイドによる三輪車の運転操作行動実験について述べる。これまでに 3.5 節や図 3.23 など提案した操作インターフェースは、人間がロボットの挙動や外界環境を観察した上で、コントローラデバイスに割り振られたコマンドを指定する操縦形式であった。この形式の場合、逐次的にヒトの目で状態を観測しながらきめ細かく操縦指令の入力を与えていく必要がある。大型障害環境における運転操作行動では、ヒトの目によるきめ細かな修正作業から独立した操縦形式を構成する。

これまでのコントローラデバイスを用いた操縦形式と、認識行動システム（図 4.17）に基づいた操縦形式の差異を簡略化して表した構成図を図 4.18 に示す。図 4.18 の左側のデバイスがジョイスティックコントローラをはじめとする操縦デバイスであり、右側が `move_base` [79] の機能を用いた計算機上の Viewer からの操縦指令である。左側のコントローラデバイスでは人間が逐次複数回の操作コマンドを入力するのに対して、右側ではモニターに表示された Viewer のみを人間が見てゴール地点を指示するだけの操縦作業となっている。ゴール指示が与えられると、作成した大型障害環境に対するコストマップに基づいて、大型障害物に干渉

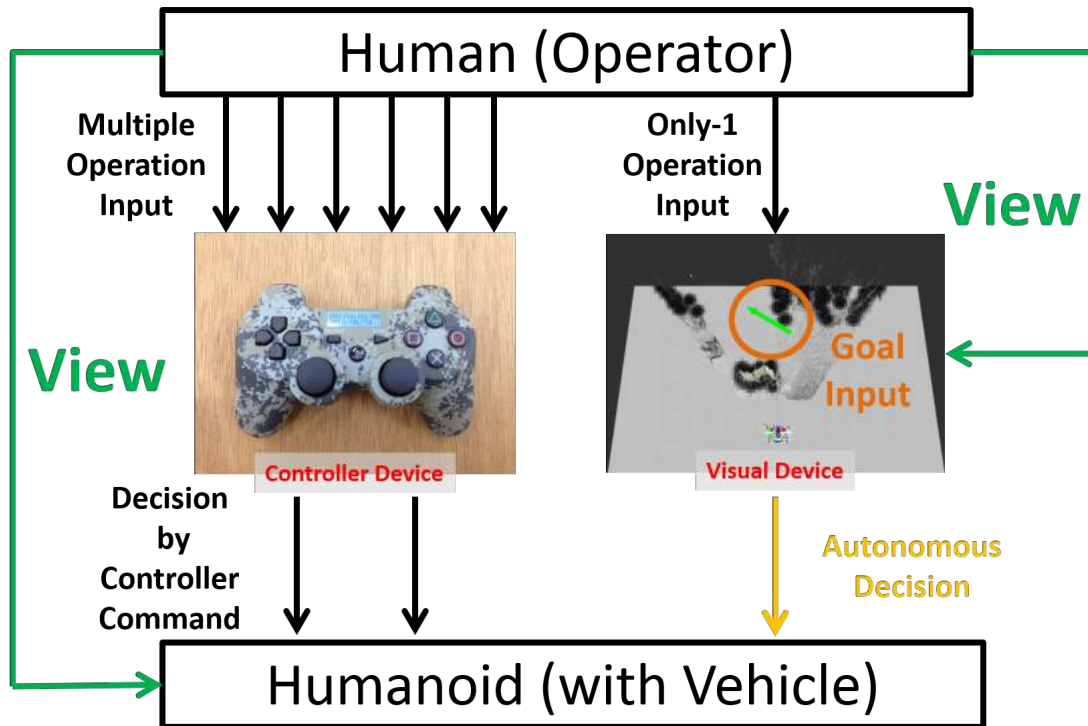


図 4.18: Difference between Controller Device and Visual Device on Operating Interface
Visual device provides the autonomous decision for vehicle manipulation by humanoid robot.

しないための機器操作行動をヒューマノイドが自律的に決定し実行する操作インターフェースとなっている。

構成した操作インターフェースを用いて、大型障害環境下で障害物に衝突しないような操作行動決定を実証した実験を図 4.19 に示す。行動実験では、ゴールの指示入力のみを ROS [140] が提供している RViz Viewer¹⁾ から指令を送っている。提示されたゴールに対して、ヒューマノイドは適切に三輪車のハンドル操舵・クランク操作の入力を決定しながら大型障害物を回避するような行動生成を実現している。

本実験の検証から、構築したシステムにおける以下の有用性が示された。

- 大型障害環境における外界環境認識法と行動に対する軌道修正法の妥当性
- 運転操作のための制御系と外界環境認識系の両システムに対する操作インターフェー

¹⁾<http://wiki.ros.org/rviz>

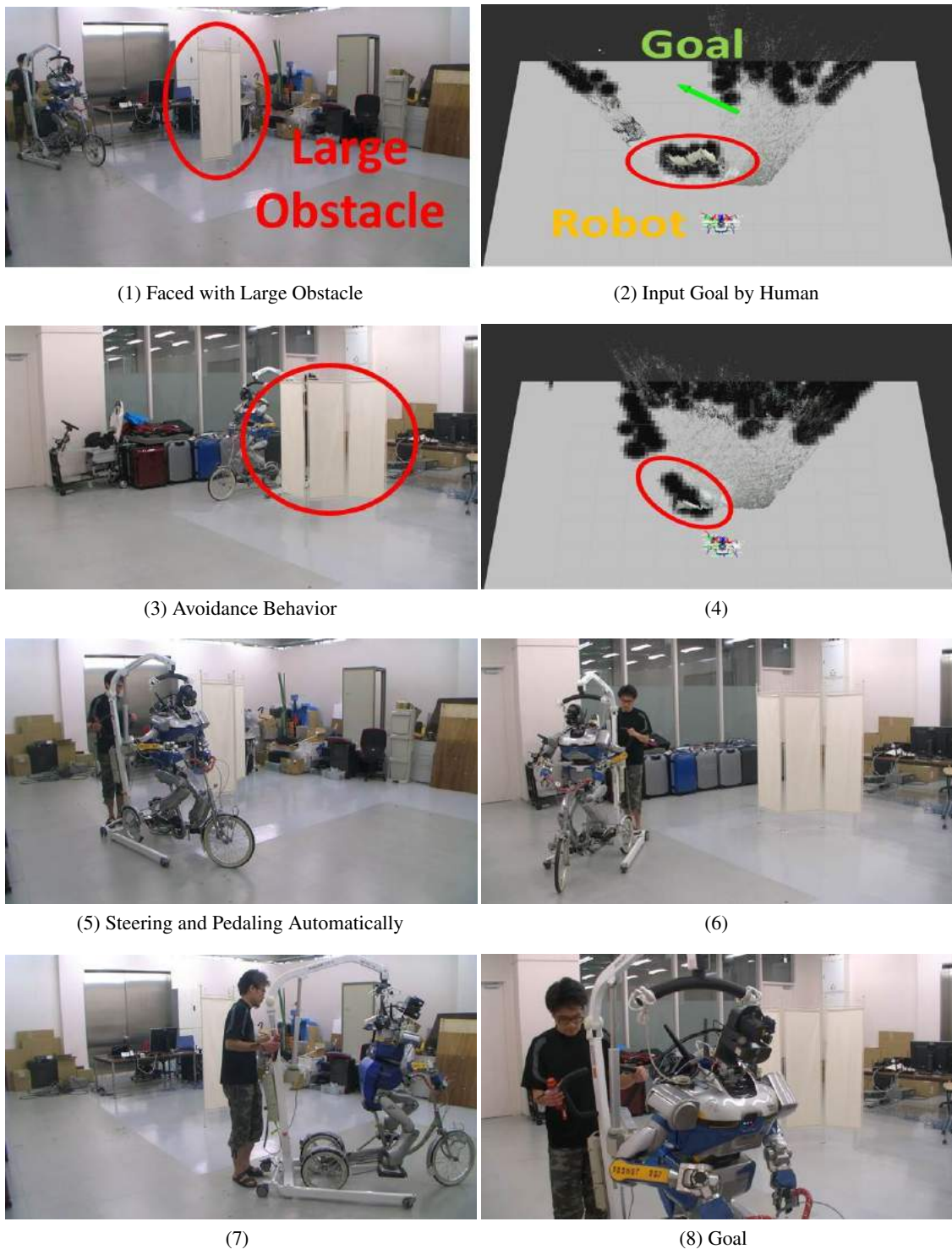


図 4.19: Recognition and Manipulation Behavior for Large Obstacle Situation

Human gives only goal marker as input on 2D costmap viewer. Humanoid autonomously drives tricycle to goal avoiding large obstacle (partition wall).

スの整合性

さらに本研究では後述する 7.3.2 項において、搭乗型機器上のヒューマノイドが自己位置を把握しながら大型障害物回避を行う並行処理型の行動タスク実現についても示していく。

4.4.3 小型障害環境に対する認識行動

小型障害物の認識法

小型障害物は大型障害物と異なり、視覚認識が困難な場合の障害物を対象とする。例えば、ヒューマノイドの足元環境の認識は難しく、視野角の不可到達性や地形などの僅かな段差障害が問題となりうる。ヒューマノイドによる搭乗型機器の運転操作においては、段差地形の踏破などの課題を解く必要があり、段差障害の検知・把握ならびにそれに基づいた障害の乗り越え行動生成などを目的としていく。

本研究では、三輪車の運転操作時(漕力作用時)において外乱となる障害を、ヒューマノイドが提供する漕力の変化(高まり)から力覚的に検知する手法(Pedaling Power Detection Method: PPDM 図 4.20)を提案する。小型障害物を乗り越えるためには、搭乗型機器である三輪車とヒューマノイドからなる一体化した系の位置エネルギーを最初に稼ぐ必要がある。本手法ではこのエネルギーを、車輪の回転エネルギーのジェネレーターに相当するクランクに対してのヒューマノイドの漕力の仕事から力覚情報として認識する。すなわち、三輪車が小型障害物に衝突した際には、ペダルを踏み込む力が搭乗型機器およびヒューマノイドの脚部を介して負荷となり伝達される。本提案手法においては、この負荷をヒューマノイドの各足裏に配置された6軸力センサを利用した力覚情報から得る。

小型障害環境における行動生成法

前述の PPDM により検知した小型障害物との衝突をフラグとして得る。衝突判定後の行動生成として、乗り越え動作を目的とした踏み込み姿勢への遷移をおこなう。踏み込み姿勢遷移は、脚部の能動機能と受動機能のオフライン探索指標/オンライン探索指標を逆方向に作用させたクランクの逆回し操作により実現する。クランクの逆回転は漕力が非提供のモードであることから、小型障害物が前に存在する場合において有効な手法である。このような補正行動は、搭乗型機器を操作する際の障害衝突時の反射行動としても観察することができる [109]。クランクの逆回し操作への遷移では、通常の運転時のペダリング操作においても

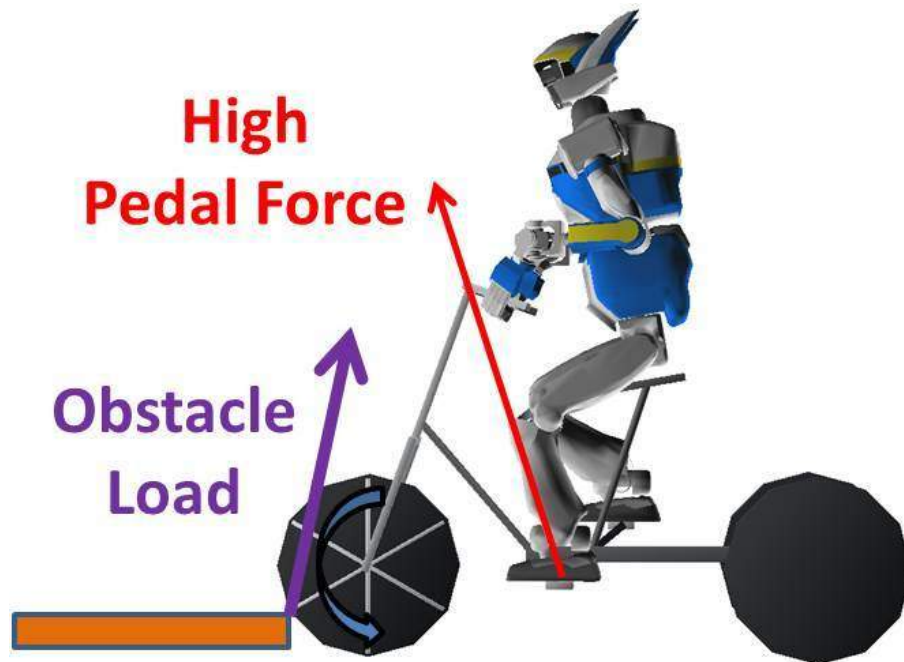


図 4.20: Pedaling Power Detection Method: PPDM [53]

Small obstacle load is detected through change of pedaling force by 6-axis force sensor on foot.

適用されている能動機能と受動機能のマニピュレータに対する探索指標の切替方策（3.4.4項）を活用している。衝突時のクランク角から切替角に到達するまでのクランクの逆回し操作による踏み込み姿勢への補正を実行することで、任意のクランク角度で小型障害物に衝突した場合でも、クランクの周期性から一意的な踏み込み姿勢への復元がなされる。

踏み込み姿勢への遷移後は、双脚の能動機能と受動機能の各探索指標を再度切り替える。ヒューマノイドの総重心を能動機能側の足裏-ペダル間の上方へと移動させ、漕力を大きく提供することで小型障害環境の乗り越え動作が生成される。

小型障害環境での軌道修正のアプローチ

認識法で述べた通り、小型障害環境では障害物を視覚情報でなく力覚情報により認識検知する。ここでの力覚情報は、大型障害環境の項で述べた視覚に基づくナビゲーション情報に相当する環境認識のパラメータであり、表 4.6 で整理したオンライン探索指標および軌道修正のパラメータの対応から、力覚情報は探索変数として捉えることができる。また、この場合に更新（修正）される軌道に相当するパラメータは、先に述べたクランクの逆回し操作の

議論より，大型障害環境の場合と類似した目標クラック角度 $\theta_c^{\text{ref}}(t)$ と考えることができる．以上の対応をまとめると，小型障害環境での軌道修正におけるパラメータは次に示す通りである．

表 4.7: 力覚ナビゲーション情報に基づくオンライン探索指標および軌道修正のパラメータの対応

探索変数	探索目標	更新される軌道
漕力の力覚情報 $f^{\text{msr}}(t)$	通常の漕力の目標指令 $f^{\text{trg}}(t)$	目標クラック角度 $\theta_c^{\text{ref}}(t)$

探索目標は，通常の小型障害物の干渉がないときの漕力の目標指令値であり，第3章で述べた $f^{\text{trg}}(t)$ に相当する．表 4.7 は，搭乗型機器の操作に対する軌道修正法として提案した受動機能の制御のオンライン探索指標に関する対応（表 3.3）に似た構成を呈しており，更新される軌道の対象が表 4.6 に対応した形式となっている．このことから小型障害環境での軌道修正法は，力覚的な操作と外界環境認識を組み合わせた構成法であることが分かる．

小型障害環境に対する認識行動システム

PPDM をベースとした小型障害環境の認識および乗り越え動作生成から構成される小型障害環境認識行動システムを図 4.21 に表す．

小型障害環境認識行動システムにおける一連のアルゴリズムを以下に示す．

1. ヒューマノイド足部に搭載された 6 軸力センサにより双脚のペダリング漕力の負荷を検知．
2. PPDM(Pedaling Power Detection Method) に基づく小型障害物の認識．
3. 双脚の能動機能と受動機能の探索指標切替操作による目標クラック角度の軌道修正．
4. ペダルの踏み込み漕力の増大を利用した小型障害物の乗り越え動作生成．
5. 手順 1 へ帰還．

小型障害物の乗り越え行動実験

図 4.22 にヒューマノイドによる小型障害環境における三輪車の乗り越え行動実験を示す．図 4.22 -(2) で外界に存在する小型障害環境(本行動実験では，14[mm]厚の木板の段差)に対

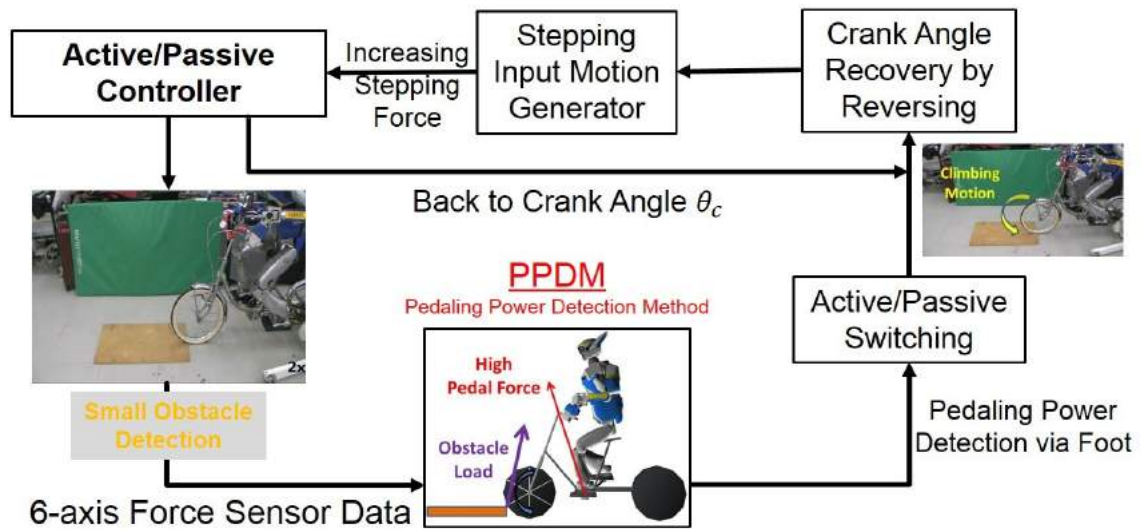


図 4.21: Detection and Planning System for Small Obstacles

して、漕ぎ進めたときの漕力による障害検知を6軸力センサをベースとした力覚情報に基づきPPDMから認識する。認識後、図4.22-(3)ではクランクが踏み込みやすい角度となるように軌道修正を行い、漕力を高めた乗り上げ動作を生成することで図4.22-(4)に示すように小型障害物の乗り越え行動を実現した。

4.5 本章のまとめ

本章では、ヒューマノイドによる搭乗型機器の運転操作行動に付随する

- 運転操作時のヒューマノイドの搭乗姿勢の診断、およびそれに基づくヒューマノイドの自己姿勢に対する軌道修正戦略
- ヒューマノイドと搭乗型機器が併合した系外の外界環境の認識、およびそれに基づくヒューマノイドと搭乗型機器への軌道修正戦略

を提案し、自己と外界に対する行動実現法を論じた。基盤となる搭乗型機器の操作に付帯するバランスの修正や外界に存在する外乱に対しての処理行動は、統合的な運転操作行動を実現するシステムを確立する上で、必要不可欠な構成要素となる。人間が必要とする感覚情報

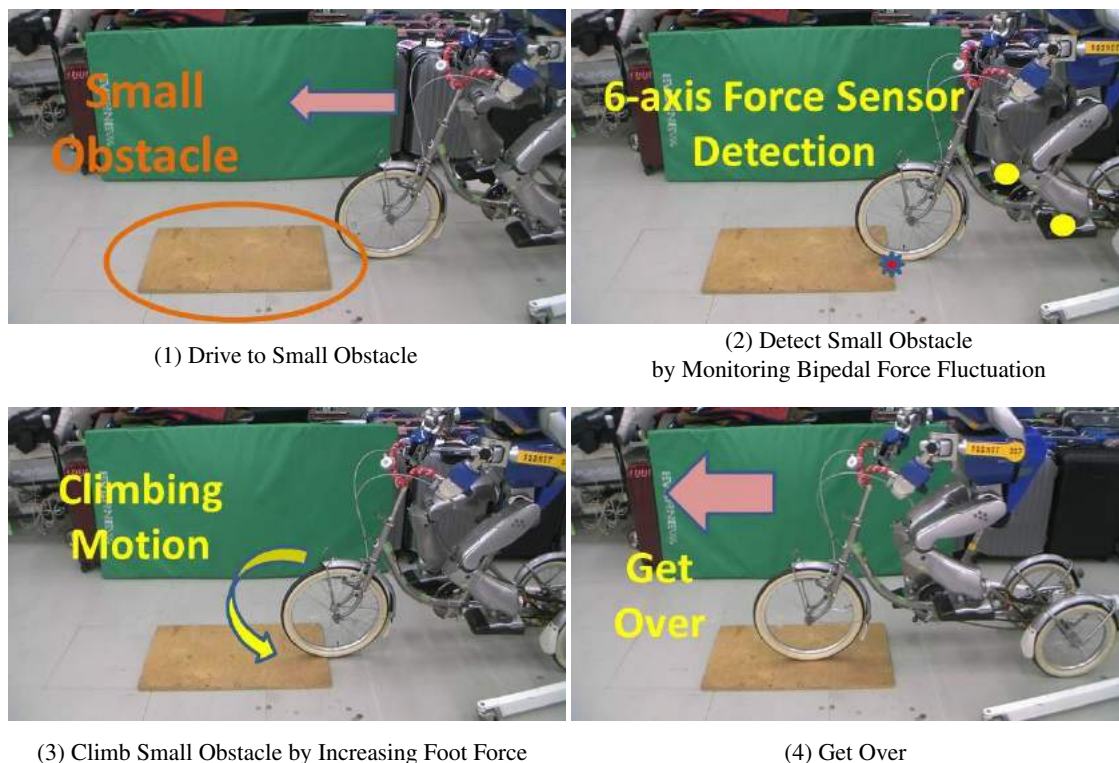


図 4.22: Recognition and Passing Behavior for Small Obstacle [53]
14[mm] thick wooden board is used as small obstacle in this experiment.

処理をヒューマノイドにも適用して実装することで、ヒューマノイドの自己身体把握と操作対象となる搭乗型機器外の他環境把握の2つの側面からの行動アプローチが実現可能となる。

また、オンライン探索指標の基である探索変数に対する認知把握処理では、ヒューマノイドロボットに搭載された単一のセンサインターフェースに拘ることのない、複合的なセンサデバイスシステムに基づいた軌道修正法を確立した。本研究ではヒューマノイドによる探索変数の感覚検知システムとして、

- 体内（胸部）に内蔵された IMU ユニットによる自己の傾斜状態/姿勢状態の診断系
- 頭部に取り付けられたステレオカメラによる環境認識点群データのオンライン取得に基づく外界の大型障害環境に対する認識系
- 足部に搭載された6軸力覚センサによる検知をベースとした視覚認識困難な外界の小型障害環境に対する認識系

を構築した。上記に示したようなロボット自身が保有するセンサインターフェースの多様性は、ヒューマノイドロボットがもつ大きな貢献性の一つである。

第5章

搭乗する環境に適応したゲイン学習

5.1 はじめに

これまでの章では、第2章で述べたオンライン探索指標の概念の内、オンライン制御指標を利用した軌道修正法について主に論じてきた。本章では、オンライン探索指標のもう一つの考えであるオンライン学習指標を軸にしたゲイン学習法を述べる。特に、ヒューマノイドが搭乗型機器に搭乗する際にはヒューマノイドが属する環境の系が変わることへの対応が必要不可欠であり、我々人間が観察の上で調整を行っている制御系のゲインチューニングをロボットが自律的に学習して更新する適応行動が求められる。本章はヒューマノイドによる乗り込み動作を行動対象として、ゲイン学習法の確立と適応行動の実機検証を目的とする。

5.2節では、従来の歩行行動との違いにも着目しながら、搭乗する環境への適応行動において学習の対象となるヒューマノイドに実装された制御系について詳しく述べる。5.3節では、5.2節で述べた制御系のゲインに対する具体的な学習方策としてオンライン学習制御の提案手法を詳説する。5.4節では、5.3節で提案したオンライン学習制御法に基づいた等身大ヒューマノイドロボットによる平行二輪電動スクーターへの乗り込み行動実験を実証し、提案手法による貢献に関して比較・考察を示す。

5.2 ヒューマノイドにおける足部トルクダンピング制御系

ヒューマノイドが搭乗型機器の環境に適応するためには、ヒューマノイドに実装された制御系のゲイン学習（ゲインチューニング・値のセットを含む）が要求される。この必要性に関しては2.2.2項で詳しく述べた。本節では、2.2.2項で論じた必要性を背景に、具体的にヒューマノイドの制御構成におけるどの制御系に対してゲイン学習を実施するかについて論じる。はじめに、これまでに提案されてきたヒューマノイドの歩行におけるロコモーション（“歩行系”）と、搭乗することにより環境が変化することへの適応（“搭乗環境適応系”）の特性の違いについて議論する。議論の内容を踏まえた上で、違いにより生じる搭乗する環境への遷移における問題を観察・提起し、適応的にゲインが学習されるべき制御系について述べる。

5.2.1 歩行系と搭乗環境適応系の違い

ヒューマノイドを代表とする二足歩行ロボットの最大の目的および得られる成果は、様々な環境上での移動を可能とする歩行によるロコモーションである。二足歩行ロボットが満た

すべき必要最低限の行動生成として、まずは標準的な床面上での歩行生成がベースとなり、そこから派生・発展した多様な地形への対処が主な研究・開発方針となる。先行研究として、平坦な床面上での歩行の安定化制御法 [55] を基礎に、傾斜を含んだ不整地上の歩行 [160] や段差上への乗り上げ歩行 [16] 等、幅広い地形環境への歩行動作の適応行動が取り組まれてきている。これらの行動例においては、ロボットが歩行する環境は舗装路面や凹凸のある不整地、高さが変化する段差など様々であるが、共通する特性として地上の環境に固定された地形である点が観察できる。本研究では、このように環境に固定された地形上での歩行行動を“歩行系”（図 5.1 左側）と定義する。歩行系が対象とする歩行環境は、地形の幾何的形狀に変化がみられるが地球表面に固定されていることから動的に地形の運動が変化するような特性はもたない。

これに対して、本研究が目的とする搭乗型機器への適応行動では、図 5.1 の右側に示すように足部が接する地形において動的に変化する運動を伴い、不規則に可動する構造を特性にもっている。このような搭乗型機器上への適応行動を“搭乗環境適応系”（図 5.1 右側）として定義する。

従来の歩行系における安定化制御法では、文献 [55][56][57] をはじめ多くの手法が提案されている。従来手法の大きな枠組みとして、多様な地形形状の変化に対応した歩行の安定化制御の新規提案や改良によるアプローチがとられている。提案・実装した制御系を基に、歩行路面の傾斜・勾配の度合いや高低差のある段差など異なる環境上でのヒューマノイドの振る舞いを観察しながら、人間が制御系の調整（ゲインチューニング）を行い、提案した制御則の妥当性を検証することが多い。

一方で、搭乗環境適応系における安定化制御では、搭乗型機器の複雑かつ不規則な運動や可動的な構造により逐次的に人間が状況に応じてゲイン調整を行うことが困難である。ヒューマノイドが車の運転タスクと運転後に車から降りるタスクを継続して行う行動例 [38] では、それぞれのタスクにおける行動戦略を独立して設計した上で、各々に対する制御系の調整を行っているが、調整時におけるヒューマノイドの基本姿勢は車の座面に着座した搭乗姿勢であり、また車のロールケージをヒューマノイドが把持することでバランスを補助した安定的な状態である。これに対して、両足でのペダリング操作を大きく含んだ三輪車操作 [52][53] や、足裏自体が移動体となるアイススケート・スケートボードでの移動行動 [161][162][163] など、バランス安定化のための制御系を導入した研究も進みつつあるが、実験的なゲイン調整によりバランスを崩しヒューマノイド実機が転倒するなどの試行実験の煩雑化が問題と

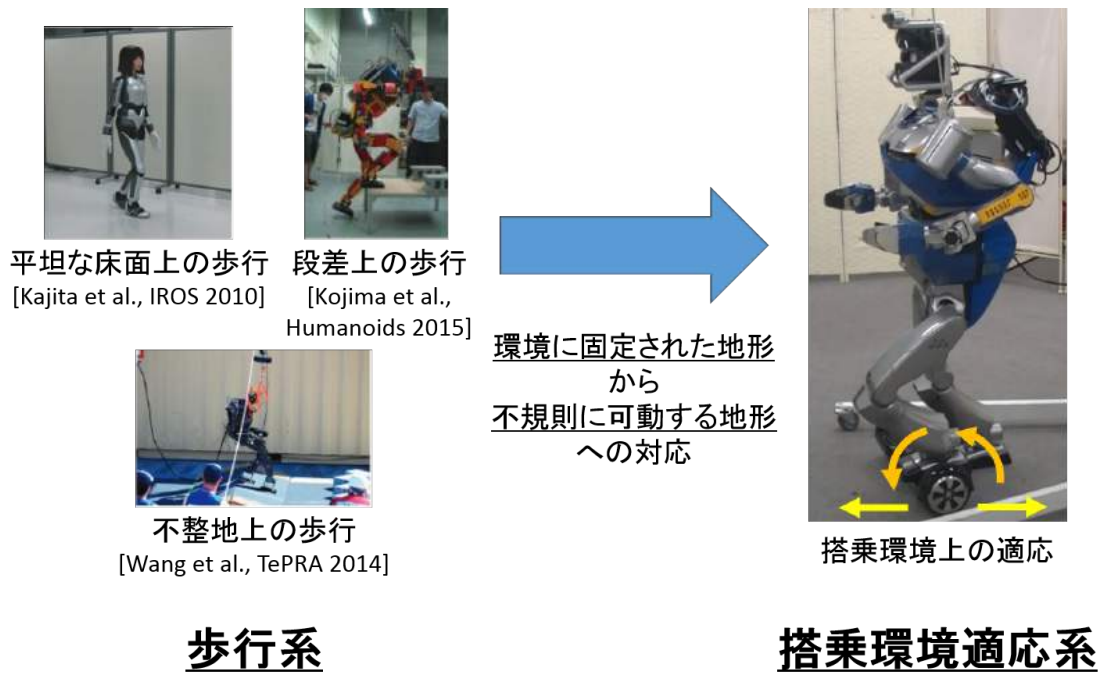


図 5.1: “歩行系” と “搭乗環境適応系”

“歩行系”においては地表の環境に固定された地形を対象とし，“搭乗環境適応系”においては搭乗型機器の可動的な構造により動的な変化を伴う地形を対象とする。

なっている。

以上の背景から，本研究の搭乗環境適応系における安定化戦略として，独自の安定化制御を開発した上でヒューマノイドに新規実装し試行的に制御系調整を行うアプローチとは異なる視点で，既存の安定化制御系に対してゲインを自律的に調整・チューニングする学習制御のアプローチを新たに提案することでこの問題を解決する．歩行系における安定化のアプローチと，搭乗環境適応系における安定化のアプローチの違いを整理すると以下の通りである。

- 歩行系における安定化のアプローチ
 - ⇒ 地形の幾何的形狀（傾斜や勾配，段差等の高低変化）に対応した安定化制御手法の提案・実装（制御系のゲインは人間が調整）
- 搭乗環境適応系における安定化のアプローチ
 - ⇒ 地形の動的変化（運動，可動的な構造の変化）に対応した安定化制御系の適応的な

ゲイン学習

5.2.2 足部トルクダンピング制御系

搭乗環境適応系に対する安定化戦略として、既存で導入されている安定化制御系のゲイン学習による適応的なアプローチの考えを前項で議論した。本項では具体的にゲイン学習の対象とする制御系について述べ、その制御則についても示す。

はじめに、歩行系において安定的となるパラメータ（制御系のゲイン）を与えた場合のヒューマノイドによる搭乗型機器への乗り込み動作を観察する。図 5.2 は等身大ヒューマノイドロボット HRP2-JSK [110] による平行二輪電動スクーター [58] への乗り込み動作を実行したときの行動実験のキャプチャであり、最後のキャプチャ（図 5.2-(8)）に示すようにヒューマノイドロボット実機は転倒した。転倒前の状態を観察すると、ヒューマノイドの両足部がスクーターのボード上に着地する前に不安定となっており、片足（左足）がスクーター上に接地し支持期となった状態から問題が生じていることが観察できる。すなわち、ヒューマノイドの足部が接する地形に関して、平坦な床面から動的に変化する搭乗型機器の上面へと遷移するタイミングで不安定となり、バランスを崩していることが実験のキャプチャより判断できる。本実験における足部まわりの状態に関する定量的観察を 5.4 節で後述する。

以上の観察から、地形に最初に環境接触するヒューマノイドの身体部位に相当する足部に対しての適応的な制御系のゲイン学習が要求されることが推察される。本実験で導入している既存の安定化制御系として、Kajita らが提案したバランス安定化制御 [55] の手法を実装したソフトウェアシステム¹⁾が、OpenRTM [164] 上の RT-Component の方式としてリアルタイム実行処理されている。この処理系の内部における足部に関する制御系のレイヤーは、文献 [55] で提案されている“Foot torque control”と“Foot force difference control”の2つのコンポーネントであるが、“Foot force difference control”は左右の両足部が接地状態にあるときの左右の足裏反力の差分に対する分配器であり、両足支持期において有効な制御系となっている。このことから、本研究では文献 [55] の“Foot torque control”をゲイン学習の対象とする安定化制御系に設定し、“Foot torque control”の制御系のゲインに対するオンライン学習を目的とする。

¹⁾<https://github.com/fkanehiro/hrpsys-base>

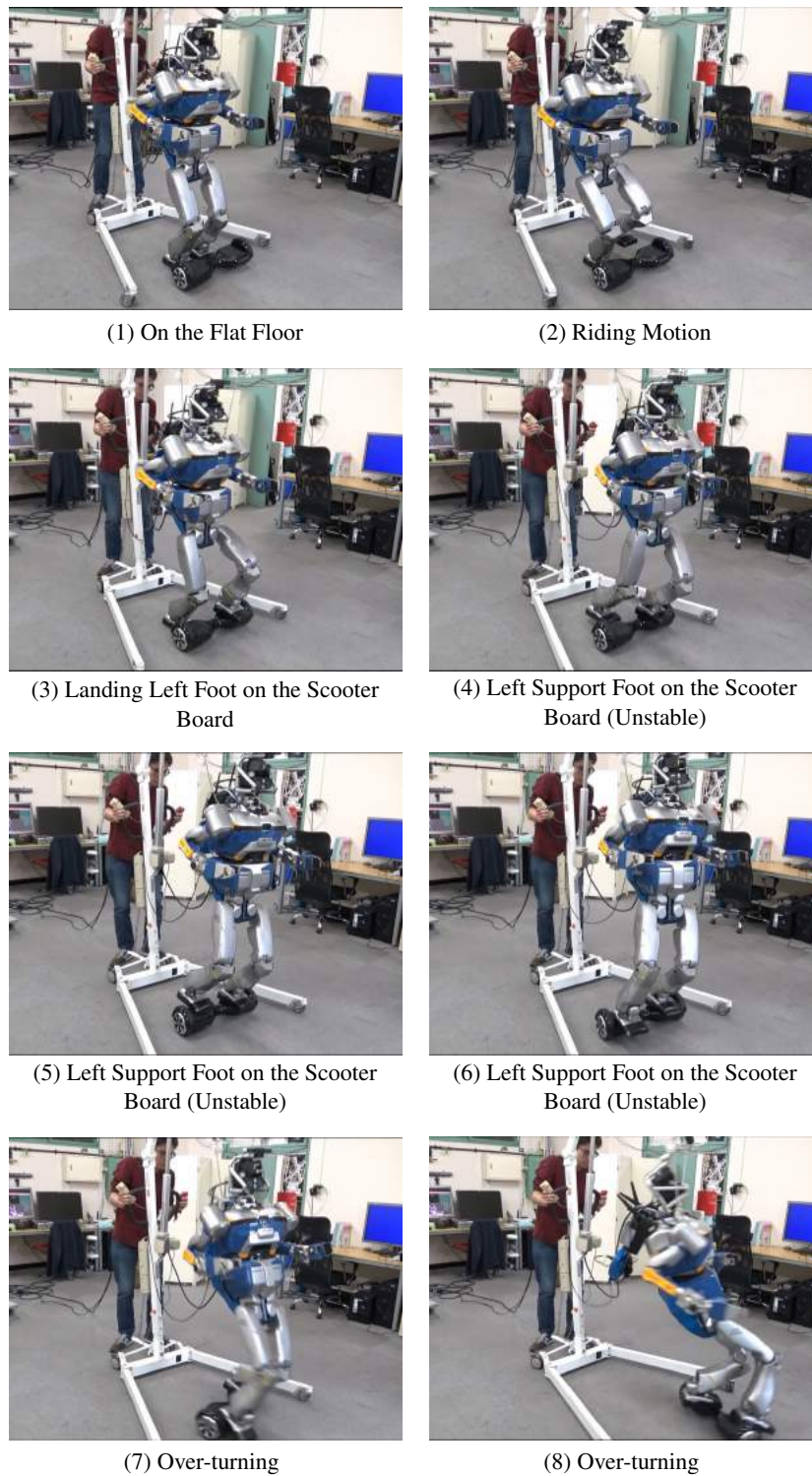


図 5.2: Failed Experiment of Riding into Parallel Two-wheeled Electric Scooter [58] by Life-Sized Humanoid Robot HRP2-JSK [110]
In this experiment, we apply the stable parameters (controller gains) on the flat floor.

文献 [55] で Kajita らが提案している “Foot torque control” の制御式を以下に引用する。

$$\delta = D^{-1}(\tau^d - \tau) - \frac{1}{T}\delta \quad (5.1)$$

式 5.1 において、 D がゲイン学習の対象となるダンピングゲインであり、 T が時定数である。 τ が環境接触の際に足部まわりで発生するトルクであり、ヒューマノイド足裏に搭載された 6 軸力センサにより計測される制御量である。 τ^d は文献 [55] で示されている別の制御系のレイヤー “ZMP distributor” から求められた足部まわりのトルクの参照値である。そして、 δ が足部まわりの姿勢（図 4.8 で示した足部フレームの姿勢 $\mathbf{R}_R, \mathbf{R}_L$ ）における角度（Roll と Pitch）の変位を微修正する制御入力を表す。

“Foot torque control” は、式 5.1 に従って τ を τ^d に制御することを目的とした足部の角度変位を操作量とするトルク制御系であり、目標値 τ^d には先述のレイヤー “ZMP distributor” 内部の演算で使用される ZMP の情報が含まれる。制御系におけるトルクの偏差 e_τ を以下とする。

$$e_\tau = \tau^d - \tau \quad (5.2)$$

このとき、式 5.2 を式 5.1 に代入し、両辺にラプラス変換を施すことで、 e_τ を入力、 δ を出力とする伝達関数 $C_\tau(s)$ が以下で得られる。

$$\delta = C_\tau(s)e_\tau = \frac{T}{D(1+Ts)}e_\tau \quad (5.3)$$

本研究では、式 5.3 を足部トルクダンピング制御系、 $C_\tau(s)$ をダンピング制御器、 D を学習対象とするダンピングゲインとよぶ。以降の議論では、ダンピングゲイン D を自律的にチューニングするオンライン学習制御系を構成し、学習則の提案においてダンピング制御器の伝達関数 $C_\tau(s)$ の情報を用いる。

5.3 ダンピングゲインのオンライン学習制御

ヒューマノイドに実装された足部トルクダンピング制御系（式 5.3）内部のダンピングゲイン D を調整することで、搭乗型機器上への移乗（歩行系から搭乗環境適応系への遷移）に際したバランス安定化を図ることを前節で述べた。本節では、既存で導入されている足部トルクダンピング制御系に対して、新たにゲイン学習のための機能を導入したオンライン学習制御系の提案および構築を行う。本節で提案するゲイン学習の方策においても、これまでに

第3章や第4章で論じたオンライン探索指標を用いたアプローチの考えが共通に適用できることを示す。

5.3.1 SGD-based open-loop Learning Control (SLC)

ヒューマノイドによる床面上から搭乗型機器上への乗り込み動作では、足部トルクダンピング制御系のゲインを適応的に調整する必要があるが、図5.2の行動実験で示したように搭乗型機器上で片足が支持する期間は非常に短く、この短期間の乗り込みのフェーズに対して実行可能なゲイン学習法を提案することが課題となる。すなわち、計算機上で実装されている微小時間となるサンプリング時間（制御周期）毎に逐次的に適正なゲインを決定しアップデートする学習システムが求められる。

また、別の課題として、乗り込み動作においては以下に示すように乗り込みのフェーズが段階的に変化するため、足部トルクダンピング制御系により制御入力を与えられるプラントの対象のモデル化が非常に複雑な問題を有している。

Phase1: ヒューマノイドの両足（A, B）が床面上に接地。

Phase2: ヒューマノイドの片足 A が床面上に接地，もう一方の足 B は遊脚。

Phase3: ヒューマノイドの片足 A が床面上に接地，もう一方の足 B は搭乗型機器上に接地。

Phase4: ヒューマノイドの片足 B が搭乗型機器上に接地，もう一方の足 A は遊脚。

Phase5: ヒューマノイドの両足（A, B）が搭乗型機器上に接地。

特に Phase4 においては、可動する搭乗型機器上で支持脚となる足が存在する一方で遊脚側の姿勢も動的に変化するため、プラントの同定には工夫が必要となる。

これらの問題に対して本研究では以下のようなアプローチをとることで解決を図る。

課題 1 短期間での乗り込み動作に対応可能な即時的なゲイン学習

⇒ Stochastic Gradient Descent (SGD) [92][93][94] に基づくオンライン学習手法の学習制御への拡張。

課題 2 乗り込み動作においてフェーズ変化を伴うプラントへの対処

⇒ 制御系を仮想的に open-loop とした学習制御系の構築。

上記のアプローチを取り入れたオンライン学習制御法を“SGD-based open-loop Learning Control (SLC)”と定義する。SLCは、学習対象の制御系（搭乗環境への適応においては足部トルクダンピング制御系）を open-loop と仮想的にみて、実時間でゲインチューニングを実行することで、動的かつ即時的な乗り込み動作に適応可能な特徴をもつ。

一般的な学習制御における評価関数の設定の多くは、Iterative Feedback Tuning (IFT) [81] や Iterative Learning Control (ILC) [165][166] などのように有限時間でのデータをバッチ処理として積算する必要があり、リアルタイムでのゲイン更新が困難である。この問題に対して、SLCではオンラインでの機械学習手法として知られている Stochastic Gradient Descent (SGD) [92][93][94] の考えを用いて学習制御のオンライン性を高めている。SLCに基づくリアルタイムチューニングシステムは、先に述べた OpenRTM [164] 実装上の RT-Component により実行されており、実装された足部トルクダンピング制御系の制御周期と同じ周期で学習系が回っている。SLCの学習周期は、制御系の周期と同じ4[ms]となる。

足部トルクダンピング制御系におけるプラントは、入力を δ 、出力を τ とする制御対象であるが、プラントの伝達関数を得るためには乗り込み動作中のヒューマノイドと搭乗型機器の間のダイナミクスに加えて、Phase1~Phase5で示したような遊脚と支持脚の状態変化も考慮に入れた複雑なプラント同定が課題である。そこで、SLCのアプローチでは、図5.3に示すようにプラント出力からの帰還を仮想的に切断した open-loop（開ループ系）を考え、学習データに偏差 e_τ の他、制御入力 δ も含めることでプラントの情報をいわずにこの課題を解決する。

ゲイン学習の対象となる足部トルクダンピング制御系に SLC のオンライン学習制御系を導入したブロック線図が図5.3である。図5.3で示された水色のブロックが SLC のオンライン学習制御系であり、偏差 e_τ とダンピング制御系の制御入力 δ からダンピングゲイン D を学習し、制御系と同周期（4[ms]）でゲイン更新を実時間実行する構成となっている。SLCの手法は、正確性よりも即時性に優先度を与えた学習制御法であり、乗り込み動作完了までの補助的な学習制御系である。一方で、乗り込み動作後の搭乗型機器上での適応行動を実現するオンライン学習制御系（図5.3下部）は、即時性よりも正確性に重みを置き、実装された制御系と同じ closed-loop（閉ループ系）での学習制御により搭乗後の問題を解く。搭乗後における学習制御法については第6章で後述する。

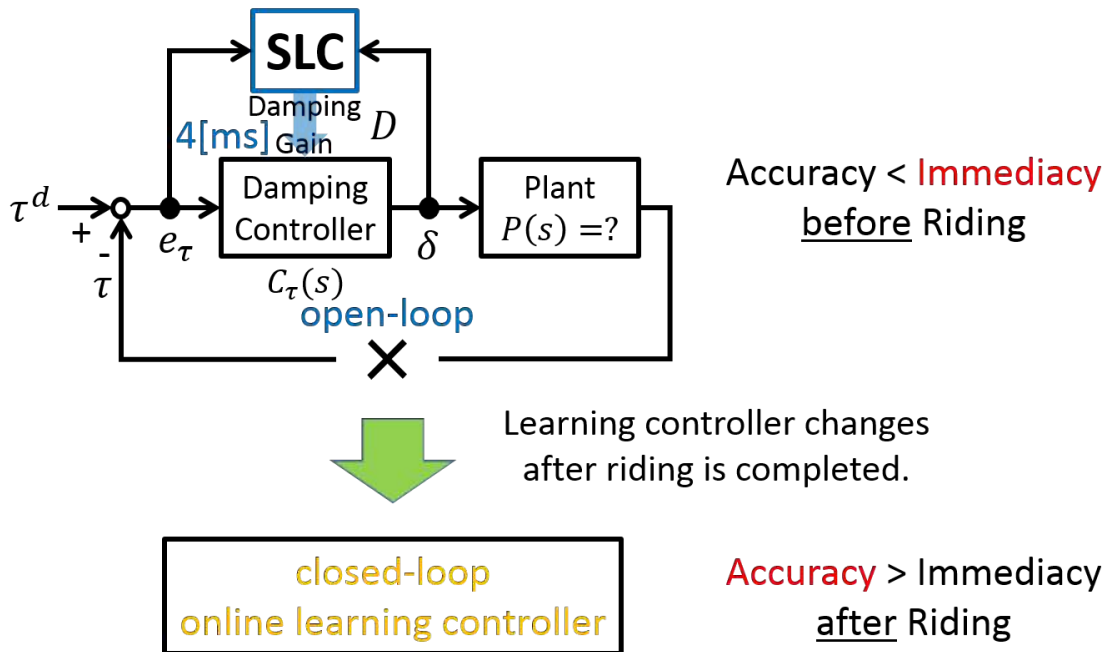


図 5.3: SLC Approach

5.3.2 SLC によるゲイン学習法

本項では、足部トルクダンピング制御系のダンピングゲイン D に関する更新則の定式化を主眼に置き、SLC の具体的なゲイン学習法のアルゴリズムについて詳説する。また、軌道修正の定式化で論じたオンライン探索指標の概念がゲイン学習においても共通に利用可能であることを示す。

評価関数の設定

SGD [92][93][94] における損失関数の設定を参考に、学習制御において最小化したい評価関数を定義する。5.3.1 項で述べたように、SLC は偏差 e_τ と制御入力 δ の足部トルクダンピング制御系における入出力状態を観測してダンピングゲイン D をチューニングする。ダンピングゲイン D を最適化するためには、トルク制御の偏差 e_τ を 0 に近付けることに加えて、制御入力 δ の発散が起こらないようにこの状態量に関しても最小化することが望まれる。そ

ここで、本研究では D に関して最小化すべき SLC の評価関数 $J_\tau(D)$ を以下の式で与える。

$$J_\tau(D) = \frac{1}{2} (Qe_\tau(t)^2 + R\delta(t)^2) \quad (5.4)$$

$J_\tau(D)$ は $e_\tau(t)$ と $\delta(t)$ の二次形式で表現されており、 Q, R は重みに相当する定係数である。以降の議論では、時間領域での状態量を表す場合は (t) を添えて表記し、ラプラス変換を施した周波数領域での状態量を表す場合は無表記とする。式 5.4 の評価関数設定では、時間 t に関するバッチ処理のデータの蓄積（積算演算や時間積分など）を含まない点が特徴であり、この特徴は SGD [92][93][94] の損失関数の設定に相当する。

学習勾配の導出

設定した評価関数 $J_\tau(D)$ の D に関する学習勾配 $\nabla J_\tau(D) = \partial J_\tau(D)/\partial D$ を導出する。はじめに、式 5.4 の $J_\tau(D)$ を D で微分することで以下の式を得る。

$$\begin{aligned} \nabla J_\tau(D) &= \frac{\partial J_\tau(D)}{\partial D} \\ &= Qe_\tau(t) \frac{\partial e_\tau(t)}{\partial D} + R\delta(t) \frac{\partial \delta(t)}{\partial D} \end{aligned} \quad (5.5)$$

次に、式 5.3 で示した $\delta = C_\tau(s)e_\tau$ の両辺を D で微分することで以下の式を得る。

$$\frac{\partial \delta}{\partial D} = \frac{\partial C_\tau(s)}{\partial D} e_\tau + C_\tau(s) \frac{\partial e_\tau}{\partial D} \quad (5.6)$$

さて、図 5.3 のように足部トルクダンピング制御系を open-loop としてみた場合、 e_τ は D の関数について定数であることから、

$$\frac{\partial e_\tau}{\partial D} = 0 \quad (5.7)$$

が成り立ち、式 5.7 を式 5.6 に代入すると $\partial \delta / \partial D$ は以下で表される。

$$\frac{\partial \delta}{\partial D} = \frac{\partial C_\tau(s)}{\partial D} e_\tau \quad (5.8)$$

同様に、仮想的な open-loop から δ を D の関数として定数とみた場合、

$$\frac{\partial \delta}{\partial D} = 0 \quad (5.9)$$

より、式 5.9 を式 5.6 に代入することで $\partial e_\tau / \partial D$ は以下で表される。

$$\frac{\partial e_\tau}{\partial D} = -\frac{1}{C_\tau(s)} \frac{\partial C_\tau(s)}{\partial D} e_\tau \quad (5.10)$$

式 5.8, 式 5.10 に式 5.3 で示した $C_\tau(s)$ を代入 ($\partial C_\tau(s)/\partial D = -C_\tau(s)/D$) して時間領域に戻すと $\partial\delta(t)/\partial D, \partial e_\tau(t)/\partial D$ はそれぞれ以下となる.

$$\frac{\partial\delta(t)}{\partial D} = -\frac{\delta(t)}{D} \quad (5.11)$$

$$\frac{\partial e_\tau(t)}{\partial D} = \frac{e_\tau(t)}{D} \quad (5.12)$$

式 5.11, 式 5.12 を式 5.5 に代入すると, SLC の学習勾配 $\nabla J_\tau(D)$ は以下で導出される.

$$\nabla J_\tau(D) = Q \frac{e_\tau(t)^2}{D} - R \frac{\delta(t)^2}{D} \quad (5.13)$$

ダンピングゲイン D の更新則

SGD [92][93][94] の手法では, 標準的な最急降下法のアルゴリズムを用いたパラメータの更新則が利用される. このことを踏まえて, 学習勾配 $\nabla J_\tau(D)$ を用いたダンピングゲイン D の更新式を以下で定める. η は学習率とする.

$$D := D - \eta \nabla J_\tau(D) \quad (5.14)$$

式 5.13 で導出された学習勾配 $\nabla J_\tau(D)$ を式 5.14 に代入する. ここで,

$$\eta Q = Q' \quad (5.15)$$

$$\eta R = R' \quad (5.16)$$

として整理することで, SLC におけるダンピングゲイン D の更新則は以下で与えられる.

$$D := D - Q' \frac{e_\tau(t)^2}{D} + R' \frac{\delta(t)^2}{D} \quad (5.17)$$

式 5.17 のように SLC では, 現在時刻 t におけるダンピングゲイン, および足部トルクダンピング制御系の入出力となる状態量 $e_\tau(t), \delta(t)$ を用いて実時間でゲイン更新が実行可能である.

オンライン探索指標に基づくゲイン学習のアプローチ

上記で導出した SLC によるゲイン学習法について, オンライン探索指標の側面から考察する. はじめに, 2.4.2 項で述べたオンライン学習指標 $f_l(\mathbf{x}(t;n))$ の定式化における評価関数 $J(\mathbf{x}(t;n))$ 内部に含まれる探索変数 $\mathbf{x}(t;n)$ を考える. $J(\mathbf{x}(t;n))$ が満たすべき条件は式 2.9 の

通りであり，条件式 2.9 を満たすような評価関数の設定として二次形式の設定例を式 2.10 に示した．SLC の評価関数として設定した式 5.4 を式 2.10 の二次形式の表現に当てはめると，探索変数 $\mathbf{x}(t; n)$ ，探索目標 \mathbf{x}_d ，および重み行列 W はそれぞれ以下の対応となる．

$$\mathbf{x}(t; n) = \begin{pmatrix} e_\tau(t) \\ \delta(t) \end{pmatrix}, \quad \mathbf{x}_d = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad W = \frac{1}{2} \begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix} \quad (5.18)$$

また， $e_\tau(t) = \tau^d(t) - \tau(t)$ であることから $\mathbf{x}(t; n) = (\tau(t) \delta(t))^T$ ， $\mathbf{x}_d = (\tau^d(t) 0)^T$ と等価である．

したがって，式 5.4 で定義した評価関数 $J_\tau(D)$ は式 2.9 の条件を満たしていることから，更新式 5.14 はオンライン探索指標に基づいたゲイン学習の一般化更新則（式 2.15）の形式として表され，ゲイン $\rho(t; n)$ および一般化更新則の重み A は以下に対応する．

$$\rho(t; n) = D, \quad A = \eta \quad (5.19)$$

以上より，SLC におけるオンライン探索指標（オンライン学習指標）とゲイン学習に関するパラメータの対応を整理すると以下となる．

表 5.1: SLC におけるオンライン探索指標およびゲイン学習のパラメータの対応

探索変数	探索目標
$(e_\tau(t) \delta(t))^T$	$(0 \ 0)^T$
更新されるゲイン	一般化更新則の重み A
D	η

$A = \eta > 0$ より A は正定値対称行列の条件を満たす．

5.3.3 SLC に基づくオンライン学習制御系

図 5.3 で示した足部トルクダンピング制御系に SLC を導入したブロック線図，および 5.3.2 項で述べたゲイン学習法のアルゴリズムを統括したオンライン学習制御系の全体構成システムを図 5.4 に示す．図 5.4 のシステムにおいて青枠の囲みが SLC に基づいた処理系であり，緑色のポートと褐色のポートで示された探索変数 $(e_\tau(t) \delta(t))^T$ の情報を基に，足部トルクダンピング制御系のダンピングゲイン D を計算機の制御周期 4[ms] で実時間更新している．各プロセスは以下の手順に従って逐次実行される．

1. 探索変数 $(e_\tau(t) \delta(t))^T$ および重み Q, R より評価関数 $J_\tau(D)$ を計算．

2. 評価関数 $J_\tau(D)$ まわりの情報と更新前の現在のダンピングゲイン D から学習勾配 $\nabla J_\tau(D)$ を導出。
3. 学習勾配 $\nabla J_\tau(D)$, 一般化更新則の重みとなる学習率 η , 更新前の現在のダンピングゲイン D , から更新後のダンピングゲインを決定し, ダンピング制御器 $C_\tau(s)$ をアップデート。
4. 4[ms] 周期で手順 1~手順 3 をリアルタイムに実行。

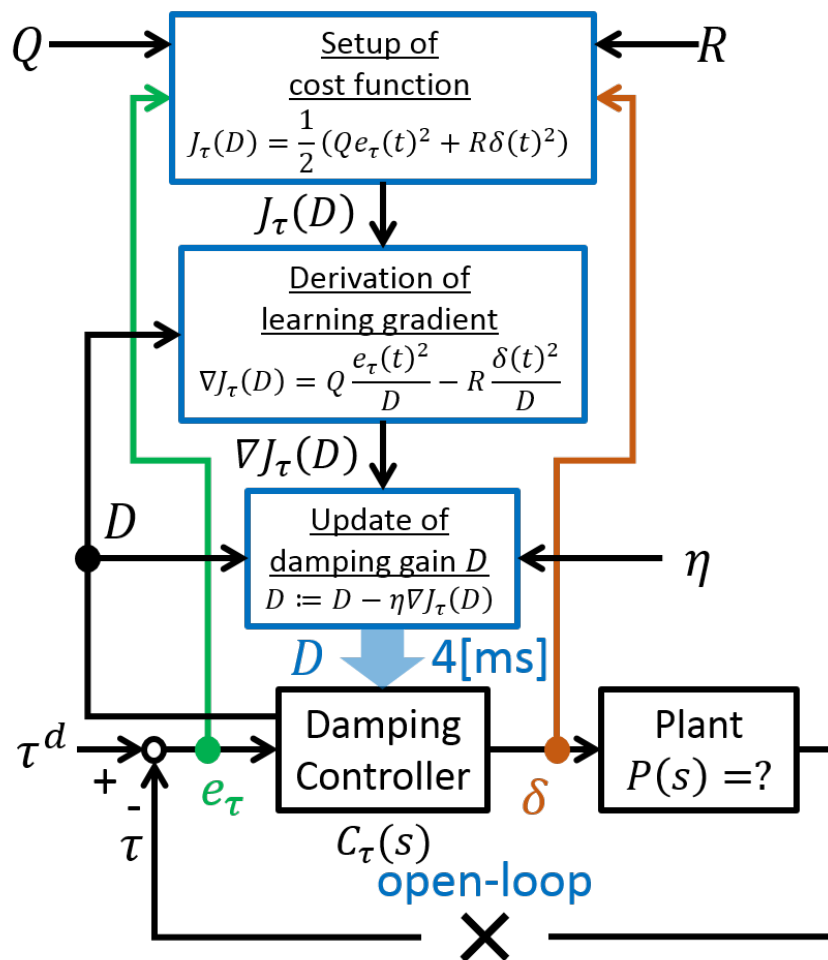


図 5.4: Online Learning Control System Based on SLC Algorithm

5.4 平行二輪電動スクーターへの乗り込み動作

本節では、5.3節で提案したSLCに基づいて、SLCの貢献を明らかにするためのヒューマノイドによる平行二輪電動スクーター[58]への乗り込み動作実験を実証する。

5.4.1 乗り込み動作におけるSLCの適用条件

SLCを適用した乗り込み動作に対する実験条件について本項で述べる。主に、以下に示す実験の条件について詳説する。

- 平行二輪電動スクーターへの乗り込み動作生成
- SLCを適用する足部トルクの対象
- SLCの実行タイミング

平行二輪電動スクーターへの乗り込み動作生成

はじめに、ヒューマノイドによる平行二輪電動スクーターに対する乗り込み動作の生成戦略を述べる。地上（平坦な床面上）でのヒューマノイドの初期立ち位置と、前方に置かれた平行二輪電動スクーターの位置関係は既知情報として、平行二輪電動スクーターのボード上に足部を乗せる際の最初の足配置（footstep [167][168]）の座標も指定されているものとする。

最初に平行二輪電動スクーターへ乗り込む足は左足として、平行二輪電動スクーターのボード上で左足が支持脚となる状態を経ながら、もう一方のボード上に右足を乗せる。このとき、右足が目標とする着地座標は、既に平行二輪電動スクーターのボード上に乗せ終えた左足からの相対座標としているため、左足が支持脚の期間で平行二輪電動スクーターが動いてボード座標が変化した場合でも、所望のボード位置に右足を乗せることが可能である。

ヒューマノイドの乗り込み動作の生成は、Nozawaらが提案しているAutoBalancer [85]により計画・実行されており、AutoBalancer [85]はKajitaらが提案した予見制御を用いたZMPの軌道パターン生成法 [84]に基づいている。乗り込み動作の軌道生成はオフラインに計画されていることから、本論文で定義したオフライン探索指標による処理として分類される。

SLC を適用する足部トルクの対象

次に、ヒューマノイドに実装された足部トルクダンピング制御系において SLC を適用する足部の成分について述べる。本研究では、即時的な乗り込み動作のフェーズにおいて最もバランス維持が支配的となる左足（平行二輪電動スクーター上において支持脚）の足部トルクに着目する。また、図 A.4 に示すように平行二輪電動スクーターのボードは Pitch 方向に対して傾斜を伴う足場環境であることから、左足の Pitch 成分（ y 軸まわり）のトルク制御器に対して SLC を適用する。乗り込み動作前のダンピングゲイン D の初期値は、平坦な床面上において適応的となる数値 $D = 22.0$ に設定しており、SLC を適用しない足部の Roll 成分では常時この数値に設定している。なお、右足の Pitch 成分に対するダンピングゲインは、左足の Pitch 成分において SLC より学習・獲得したダンピングゲインを利用した値とする。

SLC の実行タイミング

最後に、平行二輪電動スクーターへの乗り込み動作の期間において SLC を実行するタイミング（時期）を定義する。乗り込み動作における、SLC の実行開始から実行終了までの順次的なフロー図を図 5.5 に示す。実行開始、実行終了ともに足部の 6 軸力センサで計測された鉛直成分（ z 方向）の足裏反力を用いて判定を行う。

平行二輪電動スクーターへの最初の足配置において、 $f_{Lz} > F_{\text{thre}}$ の条件を満たすとき、SLC は実行開始となる（図 5.5 -左）。ここで、 f_{Lz} は左足の鉛直成分（ z 方向）反力であり、 F_{thre} は判定の閾値を表す。

その後、支持脚となる左足上方にヒューマノイドの重心を移動させて、遊脚となる右足をもう一方のボード上にリーチングさせている期間において、SLC を走らせる（図 5.5 -中央）。

もう一方のボード上への右足の足配置において、 $f_{Rz} > F_{\text{thre}}$ の条件を満たすとき、SLC は実行終了となる（図 5.5 -右）。ここで、 f_{Rz} は右足の鉛直成分（ z 方向）反力である。

本稿では、 $F_{\text{thre}} = 25[\text{N}]$ として設定する。SLC が実行終了した時を乗り込み動作の終わりと定義し、同タイミングで図 5.3 の下部で示した closed-loop のオンライン学習制御系に移行する。

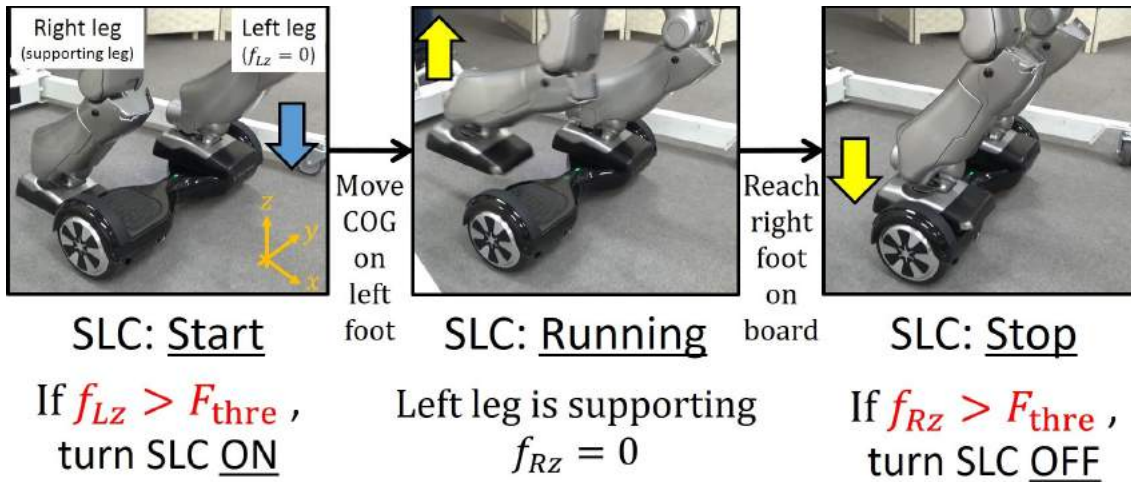


図 5.5: Sequential Flow of SLC Execution from Start to Stop

5.4.2 SLC によるダンピングゲインのオンラインチューニングの実験結果

5.4.1 項で述べた実験条件に基づき、乗り込み動作中における左足の Pitch 成分 (y 軸まわり) のトルク制御器に対して SLC を適用し、その貢献について考察する。図 5.5 で示した SLC の実行期間において、SLC を適用せずに一定のダンピングゲイン (平坦な床面上における値 $D = 22.0$) の状態で乗り込み動作を実行した場合と、SLC を適用したダンピングゲインのオンラインチューニングを実行した場合について、比較実験を行った。

等身大ヒューマノイドロボット HRP2-JSK [110] による平行二輪電動スクーターへの乗り込み動作における足元のキャプチャを図 5.6, 図 5.7 に示す。図 5.6 がダンピングゲインを一定とした状態で乗り込んだ場合、図 5.7 が SLC によりダンピングゲインをチューニングした場合をそれぞれ表す。平坦な床面で適応的にバランスを維持していたパラメータ $D = 22.0$ をそのまま用いて乗り込み動作を生成した場合、平行二輪電動スクーター上では適応的にバランスを安定化できずに図 5.6 -(7),(8) に示すように転倒した。一方で、SLC を用いてオンラインでダンピングゲインの学習を行うことで、適応的に平行二輪電動スクーター上への搭乗行動が実現された。

これらの比較実験に関して、定量的な状態観察の比較を行った結果が図 5.8 である。図 5.8 に示すプロットの時間軸は、図 5.5 で説明した SLC の実行開始時を $0[s]$ として、SLC の実行

終了時を終端の時刻としている．図 5.8 -(a1),(b1),(c1),(d1) が SLC を適用しない場合のデータであり，図 5.8 -(a2),(b2),(c2),(d2) が SLC を適用した場合のデータである．以下にそれぞれのグラフが表す状態量を述べる．

図 5.8 -(a1),(a2)： SLC の実行期間におけるダンピングゲイン D の時間変化

図 5.8 -(b1),(b2)： 文献 [55] の ZMP distributor より算出された左足の Pitch 成分の目標トルク τ^d と，計測されたトルク τ

図 5.8 -(c1),(c2)： 足部トルクダンピング制御系からの制御入力 δ

図 5.8 -(d1),(d2)： 式 5.4 より計算される評価関数の値 J_τ

式 5.3 のダンピング制御器 $C_\tau(s)$ の時定数 T ，および式 5.4 の評価関数の重み Q, R ，式 5.14 の学習率 η はそれぞれ表 5.2 に示す値に設定する．

表 5.2: 乗り込み動作実験におけるパラメータ T, Q, R, η の設定

T	Q	R	η
1.5	0.1	1.0×10^5	0.01

SLC の実行の総時間は約 1.2[s] の結果となった．このことから，即時的なオンラインでのゲインチューニングが要求されることが確認できる．SLC を適用した図 5.8 -(a2) の結果の通り，ダンピングゲイン D は初期値 $D = 22.0$ より学習を開始して，増加しながら SLC の終了時点では約 $D = 40.0$ の値を獲得した．

SLC を適用しない場合の足部トルクダンピング制御の偏差は図 5.8 -(b1) のように増加しており，これに対する制御入力 δ は図 5.8 -(c1) のように発散が観察された．また，図 5.8 -(c1) の終端時刻においては設定した入力制限に飽和し，ヒューマノイドロボット実機は平行二輪電動スクーターから転倒してしまった（図 5.6 -(7),(8)）．これらの観察事項は，図 5.8 -(d1) の評価関数の増加傾向からも明らかである．

一方で，SLC を適用した場合の足部トルクダンピング制御の偏差は図 5.8 -(b2) のように終端時刻においても残留が観察できるが，図 5.8 -(b1) と比較して段階的に増大することはなく，制御入力 δ についても図 5.8 -(c2) のように発散を防ぐことに成功している．さらに，図 5.8 -(b2) で観察される目標トルク τ^d は，図 5.8 -(b1) と比較して大きな発散はみられない．

このことは、文献 [55] の ZMP distributor 内部の計算で使用されている ZMP の項に関連付けられ、図 5.6 -(7),(8) や図 5.8 -(b1) のように不安定に発散し転倒することのない乗り込み動作を実現していることが考察できる。SLC の効果は図 5.8 -(d2) の評価関数の変化からも観察でき、1 つめのピーク (0.6[s] 付近) に比べて 2 つめのピーク (1.0[s] 付近) が小さくなっており、逐次的に最適化に向かって学習制御が進行していることが判断できる。

以上の実験結果から、足部 (本実験では左足) の接地する環境が床面上から平行二輪電動スクーター上へと遷移する際には、制御則 (式 5.3) 内部のダンピングゲイン D に関して小さい値から大きい値へと調整されることが適応的となることが裏付けられた。 D が小さい値から大きい値に修正される (ダンピング制御器 $C_r(s)$ のゲインが小さくなる) ことは、式 5.3 の足部トルクダンピング制御系の効果 (制御入力 δ) が弱められることを意味しており、床面上において支配的となっていた足部トルクダンピング制御の役割が、平行二輪電動スクーター上では平行二輪電動スクーターがもつ制御機能の一部へと分担されていることが推察される。

このように SLC によるゲイン調整法は、ヒューマノイドの足場環境が搭乗型機器上へと変化する場合に有効な適応的手法である。関連研究として、ヒューマノイドの不整地上での歩行動作において支持脚、遊脚での ZMP の移り変わりを考慮して、左右の脚に適用したインピーダンス制御の剛性を適応的に調整する脚間インピーダンス分配制御 [169][170] による手法等も提案されている。こちらの手法では左右の足にかかる鉛直成分の接触力の分配比に応じて剛性を逐次変更しており、本研究においてはダンピングゲインの逆数 D^{-1} (式 5.1) が剛性のパラメータに相当し、提案した SLC による調整法では足部のトルク制御の偏差の大きさ $e_r(t)^2$ とその制御入力の大きさ $\delta(t)^2$ の比に応じて D を調整 (式 5.17) していることに相当する。

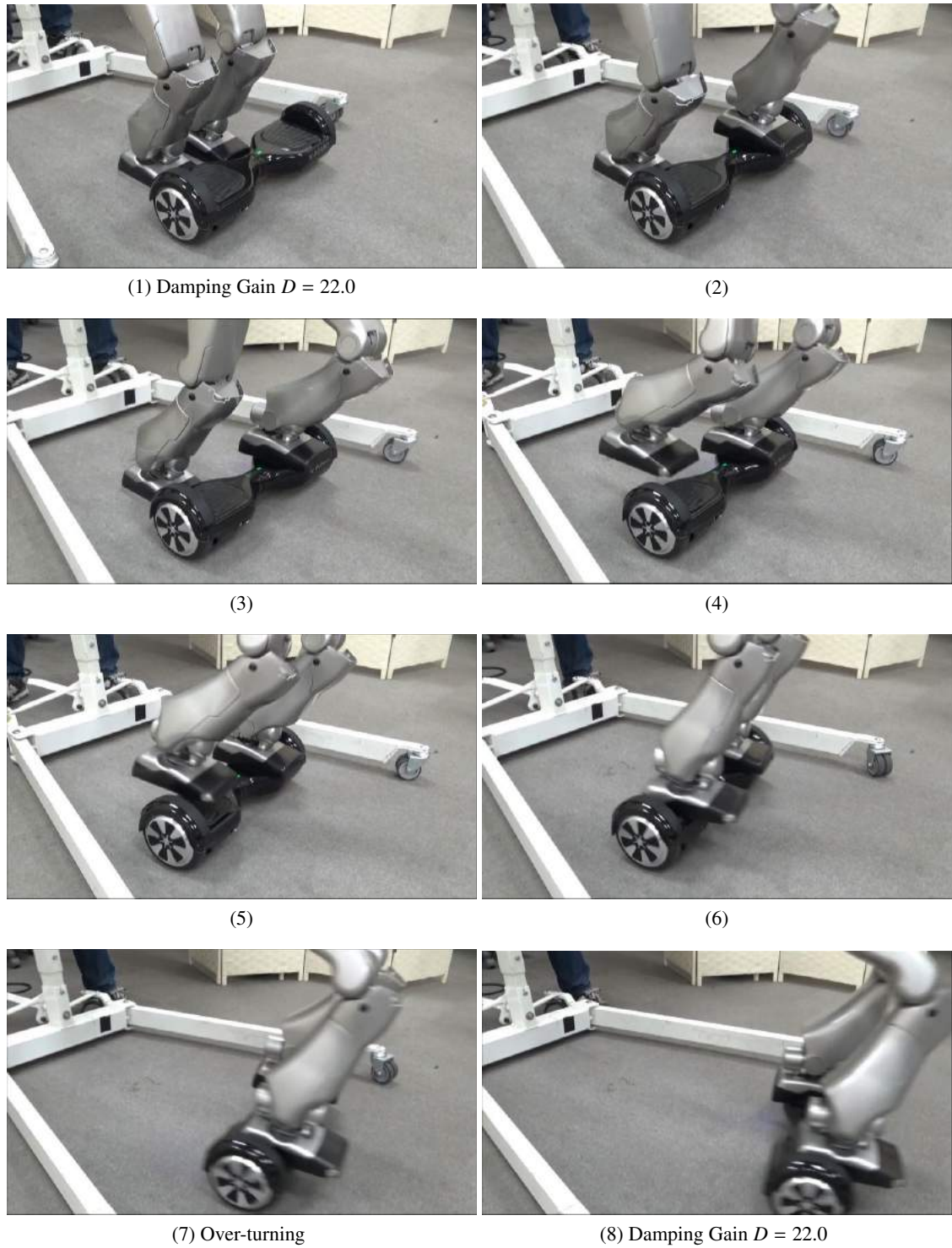


図 5.6: Riding Experiment into Parallel Two-Wheeled Electric Scooter without SLC

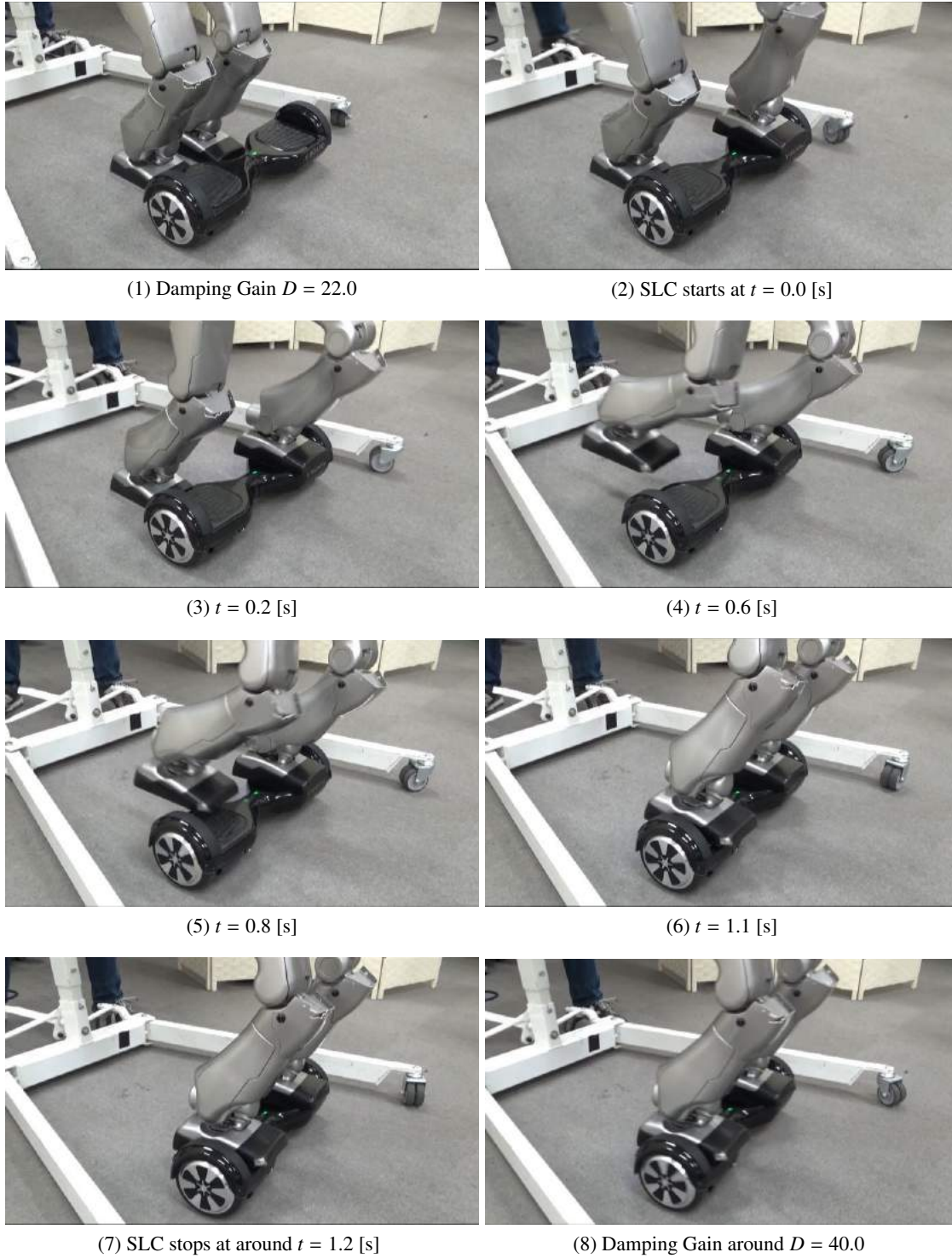


図 5.7: Riding Experiment into Parallel Two-Wheeled Electric Scooter with SLC

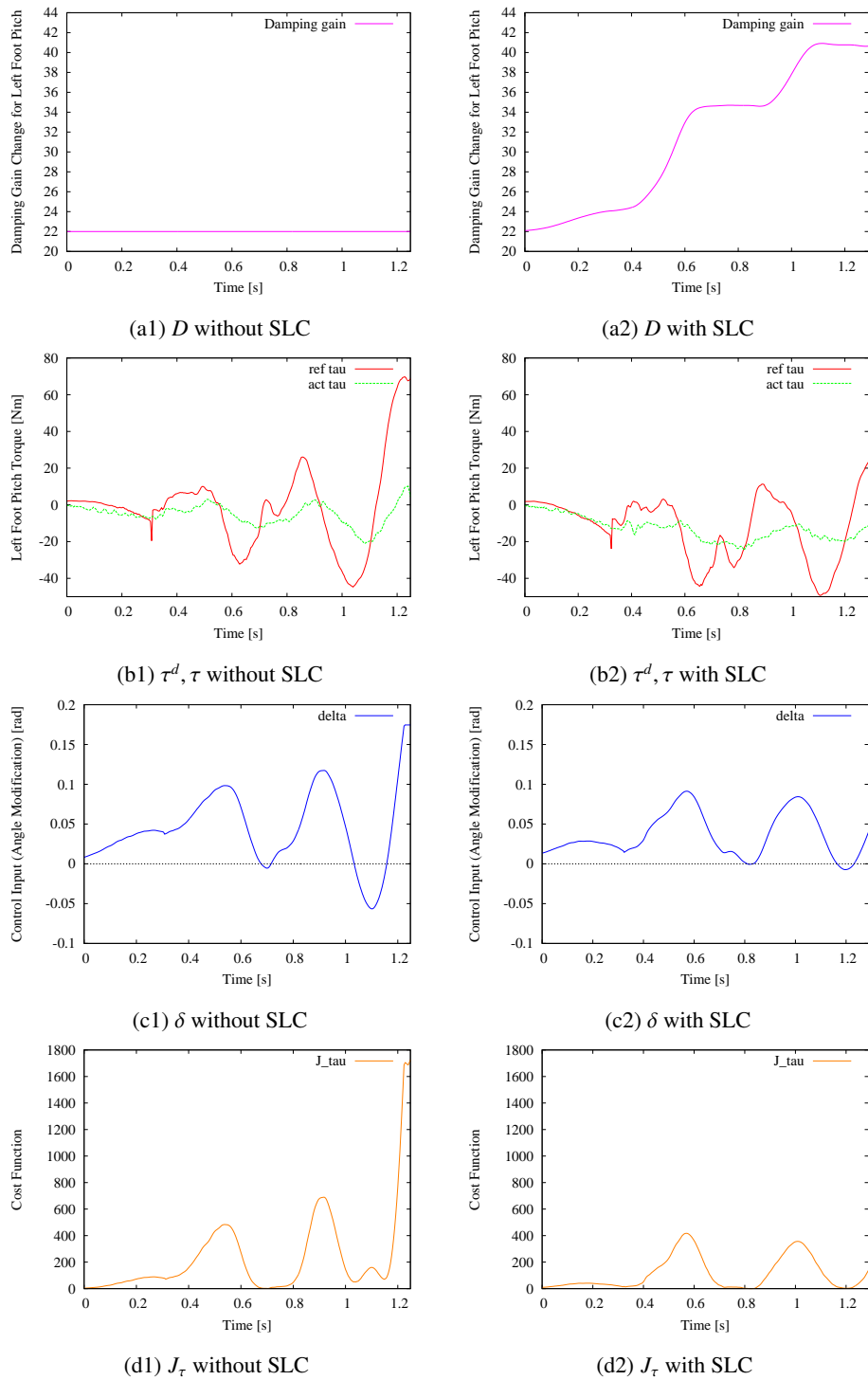


図 5.8: Comparative Experimental Results of SLC Contribution for Left Foot Pitch during Execution of Riding

5.5 本章のまとめ

本章では、ヒューマノイドにおける搭乗型機器への適応行動として、搭乗する環境（ロボットが属する地形）が変化することへの適応を目的としたゲイン学習法について論じた。ヒューマノイドによる従来の歩行行動などが対象とする環境は歩行系に分類され、歩行系では地形が世界環境に固定された条件をもつ。一方で、ヒューマノイドが搭乗型機器に乗り込む行動などが対象とする環境は搭乗環境適応系とよばれ、世界環境に固定されない不規則に可動する構造をもった地形上で相が切り替わることへの対処、具体的には転倒を防止するようなバランスの安定化を図る戦略が必要となる。本研究では、この問題に対するアプローチとして、従来の歩行系における対処法が独自の安定化制御理論の実装と人間による環境に応じた制御系の調整であるのに対して、未知な動的変化を伴う搭乗環境適応系においては、既存の実装された制御系に対して人間が干渉しない自律的な適応調整を行うことで解決するアプローチを示した。このアプローチが、本章の題目である「搭乗する環境に適応したゲイン学習」に相当している。

「搭乗する環境に適応したゲイン学習」を実世界で活用したヒューマノイドにおける行動例が、平行二輪電動スクーターへの乗り込み行動である。ヒューマノイドによる搭乗型機器への乗り込み動作は、極めて短い期間で実行される必要があるタスクであり、迅速な制御系のゲイン調整が望まれる。また、本章で示した乗り込み行動実験においては、ヒューマノイドと搭乗型機器（平行二輪電動スクーター）の幾何的な位置関係、および実装された制御則は既知情報としているものの、ヒューマノイドと搭乗型機器の間で働く物理的な特性、および実装された制御系のプロパティ（コントローラのゲイン）は未知情報である。

表 5.3: 搭乗する環境への適応行動における既知情報と未知情報

既知情報	未知情報
<ul style="list-style-type: none"> ・ヒューマノイドと搭乗型機器の間の相対的な幾何学的関係 ・実装された制御系の方式/制御則 	<ul style="list-style-type: none"> ・ヒューマノイドと搭乗型機器の間に働く物理的な特性 ・実装された制御系のプロパティ/人間が観察して調整する適切なゲイン

これらの行動課題の解決法として、ヒューマノイドが最初に地形（搭乗型機器の上面）と環境接触する部位にあたる足部の制御系「足部トルクダンピング制御系」に焦点を当て、この制御系に対するオンライン学習制御法“SGD-based open-loop Learning Control (SLC)”を提

案した。

SLCの特徴は

[特徴1] 短期間での乗り込み行動に対応可能な即時的なゲイン学習

[特徴2] 乗り込み行動における相の変化に対処したプラントレスな学習制御系

であり、[特徴1]が“SGD-based”，[特徴2]が“open-loop”の単語の意義に相当している。機械学習分野におけるオンライン学習手法SGDを学習制御分野に拡張させた成果が[特徴1]であり、仮想的に制御系を開ループとみてプラントの情報を利用しない代わりに制御器の入出力の状態からプラントレスな学習制御法へ発展させた成果が[特徴2]である。これらの成果となる特徴を整理した図を図5.9に示す。

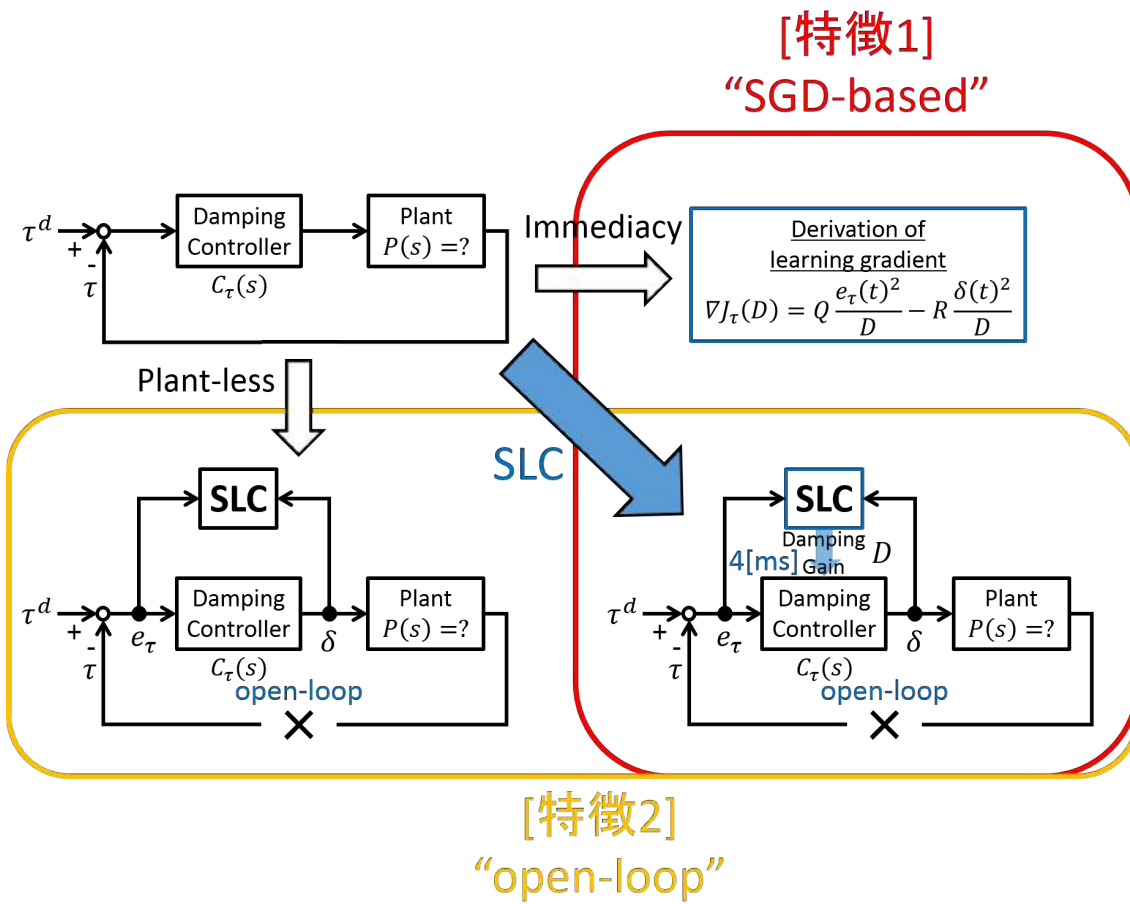


図 5.9: 本章で提案した“SGD-based open-loop Learning Control (SLC)”の2つの特徴
1つめの特徴が“SGD-based”を表し、2つめの特徴が“open-loop”を表している。

第6章

搭乗した環境の運動に適応したゲイン学習

6.1 はじめに

ゲイン学習に基づく搭乗型機器への適応行動として、第5章では搭乗する環境への移行に対する適応行動について述べた。本章では、環境移行後の搭乗した環境における運動に対する適応行動戦略をゲイン学習の側面から論じる。

6.2節では、搭乗した環境の運動をヒューマノイドが把握し、制御するための構成法について述べる。6.3節では、6.2節で提案した制御系に対してのゲイン学習方策を述べ、制御実行中にゲイン更新を適用する際の補助機能を提案する。また、これらの処理系を組み合わせたオンライン学習制御系を構成する。6.4節では、構築したオンライン学習制御系を等身大ヒューマノイドロボット実機に適用し、搭乗した環境の運動への適応を目的とした平行二輪電動スクーター上での調速行動実験を実証し提案した手法の効果を検証する。

6.2 ヒューマノイドにおける調速行動制御系

搭乗した環境においてヒューマノイドが適応的であるためには、搭乗環境に属するヒューマノイドが得る自身の運動についての自己把握が重要である。特にヒューマノイドによる搭乗型機器の運転操作を最終目標とした場合、運転による自身の移動情報としてスピード（並進速度および旋回角速度）の情報は必要不可欠である。本節では、ヒューマノイドが搭乗型機器に搭乗した後の適応行動として、運動の情報に相当するスピードに関する制御系を学習対象とした議論を展開する。本研究ではこの制御系を“調速行動制御系”と呼ぶものとする。

6.2.1 搭乗した環境の運動

“調速行動制御系”を定義するにあたって、はじめに制御対象となる搭乗した環境の運動について、運転操作を目標とした行動レベルという概念を用いて論じる。本章では、前章で取り上げた搭乗型機器として平行二輪電動スクーター [58] 上でヒューマノイドがどこまでのレベルで行動を生成すれば、所望の運動となる運転操作を実現できるかに焦点を当てる。平行二輪電動スクーターの詳細な運動の構成・構造に関しては A.1.2 項に示す。

運転操作を目標とした行動レベル

運転操作の目標と行動生成のレベルを以下のように大きく分けて定義する。

- **運転操作目標**： ヒューマノイドが転倒しないように自律的にバランスを保ちつつ、人間からの指令が与えられていない場合は移動行動をせずに待機する。
(人間から所望の速度指令を受けた場合はそれに従った行動をする.)
- **行動レベル1**： ヒューマノイドが静的に安定となる姿勢を保持した後は、平行二輪電動スクーターが実行する安定化制御のみで行動する。
- **行動レベル2**： 行動レベル1に加えて、ヒューマノイドが自身の姿勢を逐次的に修正フィードバックして安定化制御を実行しながら行動する。
- **行動レベル3**： 行動レベル2に加えて、ヒューマノイドが外界（世界座標系）から見た自身の速度を監視しながら、人間からの指令がない状況下で停止できるように速度を調整して行動する。
- **行動レベル4**： 人間が指令した所望の速度に従って行動する。

行動レベル1では、静的に安定な姿勢（ロボット関節角度）をCOG（Center Of Gravity）軌道の事前的な計算 [84] によりフィードフォワードで決定した後、安定化のための機能をすべて平行二輪電動スクーターに委ねて行動するレベルである。行動レベル2では、上記に加えてヒューマノイドロボットのLIP（Linear Inverted Pendulum）モデル化の手法 [55] を用いたトラッキング制御により、姿勢安定化のためのロボット関節角度修正のフィードバック制御を行う。

従来手法による行動レベル2までのレイヤーにおいては、平行二輪電動スクーターの安定化機能とヒューマノイドロボットの安定化制御が協調せずに独立して働いている問題が挙げられる。すなわち、平行二輪電動スクーター上から見たロボットのバランス安定性は補償されているが、平行二輪電動スクーターの外界の床面等から見た並進速度や旋回角速度は制御されておらず、予期せぬ速度で移動行動を生成する危険な状態である。

そこで本論文では、行動レベル3の行動制御法を提案することで上記の問題に対する解を与える。前提となる議論として、行動レベル1および行動レベル2の実機検証と比較・考察を交えた問題提起を次に述べる。

行動レベル1と行動レベル2の実機検証と問題点

前の議論で定義した行動レベル1および行動レベル2について、等身大ヒューマノイドロボット HRP2-JSK [110] を実際に平行二輪電動スクーター上に乗せて挙動を観察する。すなわち、平行二輪電動スクーター自体のバランス機能と従来のヒューマノイドの安定化制御法のみによる行動の振る舞いを観察する。観察事項として、二足歩行ロボットのバランス安定化の指標で広く用いられている ZMP (Zero-Moment Point) [60] の観測の他、A.1.2 項で示す移動速度との関係性を表した平行二輪電動スクーターのボード傾斜角について、ボード上で一体となるヒューマノイドの足部リンクの Pitch 姿勢角よりセンシングすることで対象の移動性を評価する。

はじめに、平行二輪電動スクーターに対するヒューマノイドの搭乗姿勢を、ボード上面にロボット足裏が乗るように文献 [84] に基づいた静的に安定な姿勢として求め、搭乗姿勢の全身の関節角度をロボット実機へ指令する。搭乗姿勢への遷移後、平行二輪電動スクーターとロボット足部を図 6.1 のような位置関係となるようにリフトから下ろして搭乗させる。搭乗後の挙動が平行二輪電動スクーターのバランス機能のみで行動する行動レベル1の振る舞いとなる。図 6.1 において平行二輪電動スクーターの中心を座標系の原点とする。この点を目標 ZMP の位置 (0,0) として、ロボット足裏に搭載の6軸力センサにより計測が可能な実 ZMP との比較を行う。

さらに、床面に固定された世界座標系 Σ_0 から見たロボット足部（右足，左足それぞれ）の姿勢変換行列 ${}^0\mathbf{R}_{\{R,L\}}$ を求めることで実時間でボード傾斜角 θ_R, θ_L を監視することが可能である。A.1.2 項で示す平行二輪電動スクーターのボード傾斜角と電動で駆動される車輪の回転速度の関係を表したプロット（図 A.6），およびボード傾斜角と搭乗者に出力されるスピードの関係式 A.4 より、並進速度と旋回角速度が0に近くなるようなオフセットの傾斜角 $\theta_{RO} = -1.7871[\text{deg}]$, $\theta_{LO} = -1.9972[\text{deg}]$ との比較を定量的に評価する。

図 6.2 は HRP2-JSK [110] の右足リンク，左足リンク，骨盤リンク，胸部リンクのそれぞれに固定された座標系 $\Sigma_R, \Sigma_L, \Sigma_B, \Sigma_C$ と世界座標系 Σ_0 との相対関係を示す図であり， ${}^1\mathbf{R}_J$ は座標系 Σ_I から見た座標系 Σ_J への姿勢変換行列を表す。 ${}^B\mathbf{R}_{\{R,L\}}, {}^B\mathbf{R}_C$ は共にロボットの関節エンコーダより既知であり，また ${}^0\mathbf{R}_C$ は胸部リンク体内に搭載された IMU (Inertial Measurement Unit) の姿勢推定により既知の情報である。

これらを用いて ${}^0R_{\{R,L\}}$ は以下の式で計算することができる。

$${}^0R_{\{R,L\}} = {}^0R_C ({}^B R_C)^{TB} R_{\{R,L\}} \quad (6.1)$$

行動レベル 2 は, 行動レベル 1 に加えて文献 [55] の LIP (Linear Inverted Pendulum) トラッキング制御を適用したロボットの姿勢に対するフィードバック修正を与える。

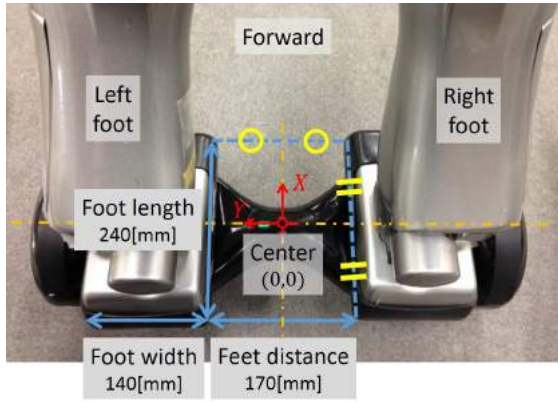


図 6.1: Feet Placement on Parallel Two-Wheeled Electric Scooter

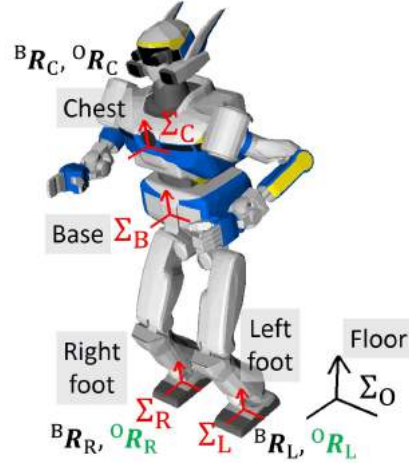


図 6.2: Relationship for Each Coordinate to Obtain the Attitude Transformation Matrix ${}^0R_R, {}^0R_L$ Viewed from the World Coordinates Σ_0 to the Right/Left Foot Coordinates Σ_R, Σ_L

行動レベル 1 および行動レベル 2 の実証実験の結果を図 6.3 および図 6.4 に示す。各々の図において, (a1)-(a4) は平行二輪電動スクーター搭乗後からの各時刻での挙動のキャプチャ, (b1) と (b2) は目標 ZMP と実際に計測された実 ZMP の時間変化の X[mm], Y[mm] のプロット, (c1) と (c2) は右足と左足の姿勢変換行列 0R_R と 0R_L から得た Pitch 成分の傾斜角 θ_R, θ_L [deg], および $\theta_{R0} = -1.7871$ [deg], $\theta_{L0} = -1.9972$ [deg] との比較結果である。平行二輪電動スクーター搭乗時の時刻を $t = 0.0$ [s] とする。

行動レベル 1 では, 図 6.3 -(b1) の時刻 0.0[s]-0.5[s] (搭乗直後) において ZMP の X 成分が瞬間的に 120[mm] を超え, 図 6.1 の両足の支持多角形から外れた不安定な状態となるが, そこから支持多角形内に戻そうとする働きが平行二輪電動スクーターのバランス機能により実現されていることが観察できる。一方で目標 ZMP とのズレは生じており, 特に図 6.3 -(b2) の Y 成分では実 ZMP が振動的に発散傾向にあることがわかる。また, 図 6.3 -(c1),(c2) の

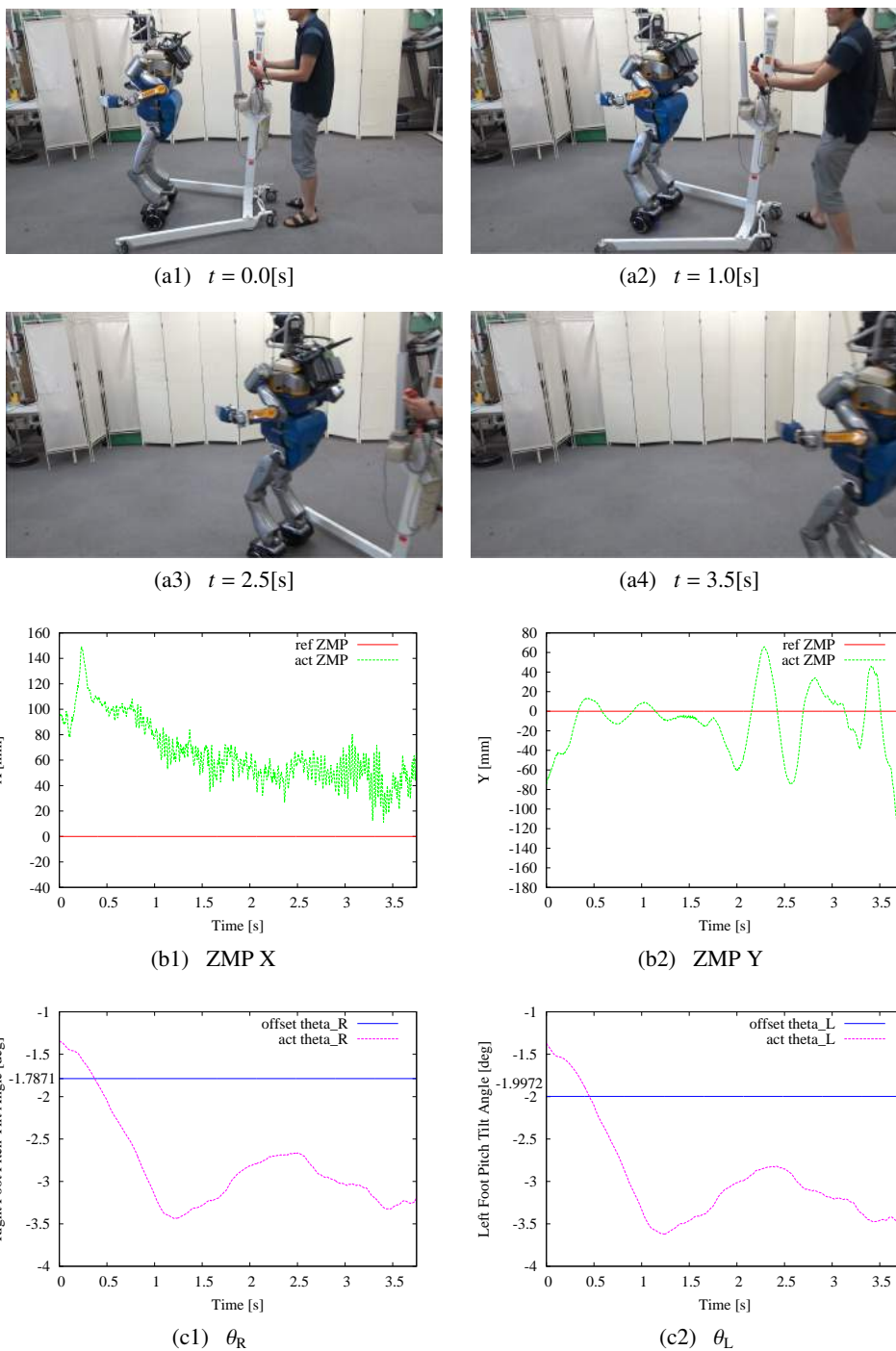
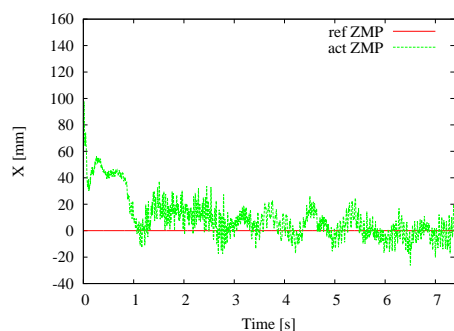
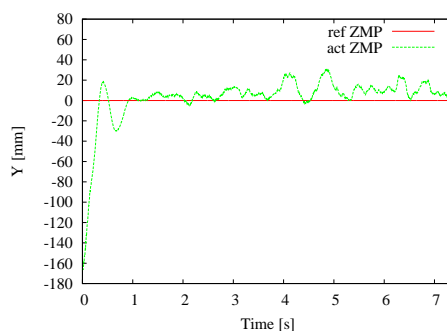


図 6.3: Experimental Results for Behavior Level 1

(a1)-(a4) show the experimental snapshots. (b1),(b2) show the plots of ZMP X and Y [mm] (the red solid lines represent reference ZMP, and the green dashed lines represent actual measured ZMP). (c1),(c2) show the plots of right/left foot pitch tilt angle θ_R and θ_L [deg] (the blue solid lines represent θ_{R0}, θ_{L0} , and the magenta dashed lines represent θ_R, θ_L).

(a1) $t = 0.0[s]$ (a2) $t = 2.5[s]$ (a3) $t = 4.5[s]$ (a4) $t = 7.0[s]$ 

(b1) ZMP X



(b2) ZMP Y

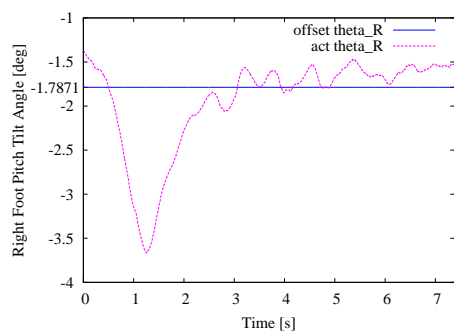
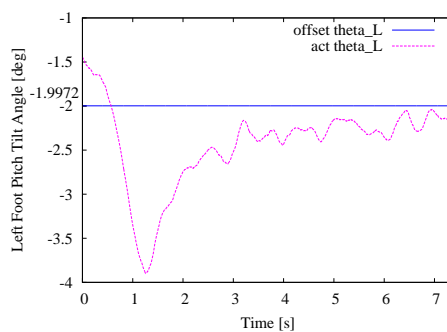
(c1) θ_R (c2) θ_L

図 6.4: Experimental Results for Behavior Level 2

(a1)-(a4) show the experimental snapshots. (b1),(b2) show the plots of ZMP X and Y [mm] (the red solid lines represent reference ZMP, and the green dashed lines represent actual measured ZMP). (c1),(c2) show the plots of right/left foot pitch tilt angle θ_R and θ_L [deg] (the blue solid lines represent θ_{RO} , θ_{LO} , and the magenta dashed lines represent θ_R , θ_L).

θ_R, θ_L は θ_{RO}, θ_{LO} から負の方向に離れていき、図 6.3 -(a1)-(a4) のキャプチャからも判断できる通り、後方に大きく加速して速度調整がなされていない状態である。

行動レベル 1 に文献 [55] の姿勢に対するフィードバック安定化制御を適用した行動レベル 2 においては、目標 ZMP と実 ZMP のズレが改善されており定常的に追従していることが観察できる（図 6.4 -(b1),(b2)）。図 6.4 -(c1),(c2) をみると、 θ_R, θ_L は θ_{RO}, θ_{LO} に対して追従的であり、一見すると速度調整に成功しているように判断できるが、図 6.3 -(a1)-(a4) のように加速的ではないにしろ後方に徐々に移動していることが図 6.4 -(a1)-(a4) のキャプチャより確認できる。

図 6.4 -(c1) は定常状態では $\theta_R > \theta_{RO}$ となっており、図 A.6 の相関関係に基づくと右車輪は前進していることになるが、実際は両輪とも後退している結果となった。これは $[\nu \omega]^T = [0 \ 0]^T$ となる平行二輪電動スクーターのボード傾斜角の目標値 $\theta_{RO} = -1.7871[\text{deg}]$, $\theta_{LO} = -1.9972[\text{deg}]$ が正確な値ではなく、これらの値は A.1.2 項で示した計測データから最小二乗法を適用して得た線形近似直線の切片であり、大まかな目安となる近似値であることが理由として推察される。また、路面の傾斜状況や平行二輪電動スクーターとロボット足裏の接触状況等によっても θ_{RO}, θ_{LO} は大きく左右されることから、この値を一定の目標値としたフィードバック制御は実用が極めて困難である。

以上の議論を踏まえた上で、平行二輪電動スクーター固有の特性（ボード傾斜角のオフセットなど）に基づいた制御法ではなく、ヒューマノイド自身が把握する搭乗環境のセンシングをベースとしたフィードバック構成を提案することで上記の問題を解決する。

調速行動（行動レベル 3）

先に述べた行動レベル 1 と行動レベル 2 で提起された課題を背景に、行動レベル 3 におけるヒューマノイド自身の世界座標系での速度把握と、その情報をフィードバックして速度を調整する行動を“調速機（Governor）”に因んで“調速行動（Speed governing behavior）”と定義する。調速行動のフィードバック制御系として、本研究では PID 制御器をベースに学習制御を適用してオンラインでゲインチューニングを行う。本研究はヒューマノイドによる平行二輪電動スクーターの運転操作を最終的な目標として定め、ヒューマノイドがオンラインで獲得したフィードバック制御器に対して、将来的に人間がゲインの最終確認・最終調整がし易いよう、人間の経験則としてチューニングが容易な PID 制御器を採用している。

フィードバック制御器に対する学習制御の手法として、反復フィードバックチューニング

(Iterative Feedback Tuning : IFT) が知られており [81][171], この手法に基づき PID ゲインを最適値に向かって逐次更新していく. IFT の処理においては, 繰り返しの学習の試行実験による実験環境設定の初期化がその都度必要となる. 初期化の作業として, PID ゲインの設定値を学習結果の値に変更する度に制御系の ON/OFF (実験のやり直し) や偏差の積分値のリセットが要求され, 等身大のヒューマノイドロボット実機を平行二輪電動スクーター上に繰り返し乗せ直すなどの煩雑かつ重労働な手間が問題となる.

上記の繰り返しの作業を無くすべく, 制御系を ON にしたまま制御実行中にその都度出力される IFT の学習結果を自動的にセットし, オンラインで PID ゲインのチューニングを行うシステムを本論文で提案する. 特に PID ゲインのうちの I ゲインを自動更新する際には, 更新前の低い I ゲイン (0 など) で修正しきれずに蓄積された過大な偏差の積分値が, 更新後の高い I ゲインで増幅され急激なオーバーシュートを伴う I 動作として働く問題が生じる. ゲイン更新の度に偏差の積分値に毎回リセットをかける単純な解決方法もあるが, 1 回の学習のサンプリング時間が短い場合は高周期でリセットがかかり I 動作として機能しなくなることや, 学習が収束する (ゲインが変化しなくなる) につれてリセットが不要になることなどが挙げられる.

制御入力に飽和制約を設けて, 過剰な I 動作等により入力が飽和した場合に I 動作を停止させたり自動整合 PID 補償器を加えたりするアンチwindアップ [172][173][174][175][176] の対策方法も知られているが, 学習制御のように学習パラメータ (PID ゲイン) の変化に対する制御挙動を観測する上では, 飽和のない正規の制御入力を与えることが望ましい. また, 調速行動における学習制御は一定の目標値への追従を目指すレギュレータ系であることから, 既存手法 [177][178][179] で提案されている目標値整形やリファレンスガバナの適用との相性も良くない.

本論文では, 制御入力に飽和制約を与えない条件のもと, PID ゲインのオンライン更新の瞬間にのみ偏差の積分量の前回値を補正することで上記の問題を解決する. 提案する補正処理の特徴として, 飽和制約の上限値・下限値の設定や多くのアンチwindアップ手法で設けられている閾値や重みパラメータなどの設定は必要なく, 更新前後の I ゲインの値のみを用いたシンプルな積分量補正を行う.

6.2.2 調速行動制御系

行動レベル 3 の調速行動を実現するためのベースとなる制御系やセンサ処理系の構成法について述べる。調速行動においては、ヒューマノイドおよび平行二輪電動スクーターからなる一体化された系の並進速度・旋回角速度のセンシングとフィードバック処理が重要となる。平行二輪電動スクーターには並進速度および旋回角速度をセンシングするためのセンサ類は搭載されておらず、搭乗者となるヒューマノイドのセンサを用いてこれらの状態量を把握する必要がある。ロボットおよび平行二輪電動スクーターの一体化した系の状態量（並進速度・旋回角速度）は、ロボットや平行二輪電動スクーターを相対とした情報（図 6.5 における Σ_{vehicle} に対する Σ_{robot} ）だけでは得られず、この系外の世界座標系から見た状態量の情報（図 6.5 における Σ_{floor} に対する Σ_{robot} ）も必要である。従来のヒューマノイドの安定化制御や全身行動生成では、ロボット相対または直接的に接触する床面や対象物を相対とする議論が多く、空中での UAV（Unmanned Aerial Vehicle）の姿勢制御などのように非接触な外界の系である世界座標系を基準とした状態量を扱うことは少ない。

以上の背景から、本論文では世界座標系における並進速度・旋回角速度の状態量の扱いを提案し、このときのフィードバック制御の考えを調速行動制御と呼ぶものとする。従来研究 [55] による行動レベル 2 のバランス安定化制御と新たに提案する調速行動制御の違いを以下にまとめる。

○ バランス安定化制御 [55]（行動レベル 2）：

ヒューマノイドが接触する環境を相対（ Σ_{vehicle} に対する Σ_{robot} ）とした状態量（CoM・ZMP など）を扱うフィードバック制御

○ 調速行動制御（行動レベル 3）：

ヒューマノイドが非接触の環境を相対（ Σ_{floor} に対する Σ_{robot} ）とした状態量（並進速度・旋回角速度など）を扱うフィードバック制御

平行二輪電動スクーター上で調速行動制御を実現するためのフィードバック制御システムを図 6.6 に示す。本構成のリアルタイム制御系は、第 5 章で述べた制御システムと同様に OpenRTM [164] 上の RT-Component の方式でロボット体内の計算機で実時間処理されており、制御周期 Δt は約 4[ms] である。図 6.6 は、次に示す 4 つの構成要素からなる PID 制御をベースとしたブロック線図であり、伝達関数として表現される。

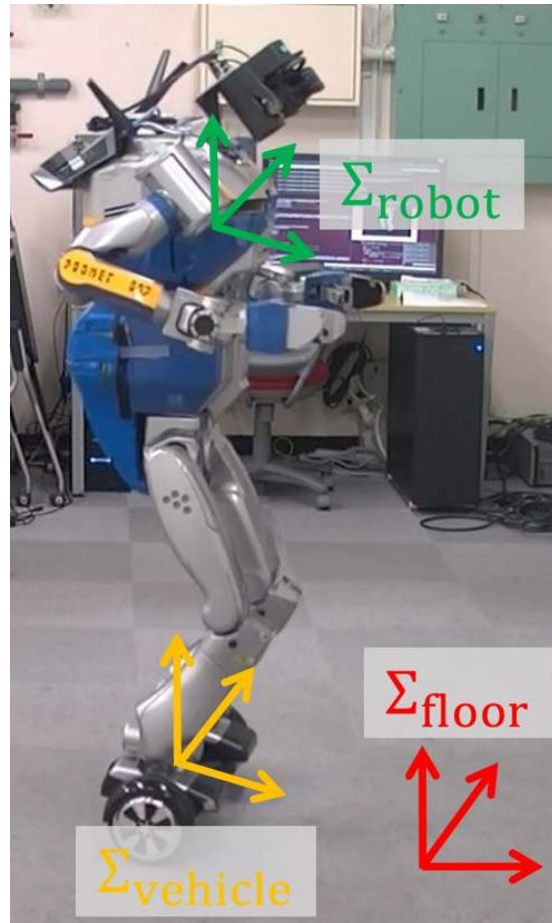


図 6.5: Relative Relationship of Coordinates Σ_{robot} , Σ_{vehicle} , and Σ_{floor}

- (A) オンライン学習制御系
- (B) 足部トルクダンピング制御系
- (C) 並進速度フィルタ
- (D) ロボットプラント

また，図 6.6 内の各パラメータの制御周期（サンプリング周期）および遅延時間（レイテンシ）を [ms] 単位で表 6.1 に示す．図 6.6，表 6.1 の各変数の説明は (A)~(D) のそれぞれの項で後述して定義する．

(A)~(D) の各構成要素について以下で説明する．

表 6.1: Sampling Period and Latency in Control System

v^d	ω^d	v	ω	e_v	e_ω
24[ms]*	24[ms]*	24[ms]*	4[ms]	4[ms]	4[ms]
u_v	u_ω	τ_{Ry}	τ_{Ly}	τ_{Ry}^d	τ_{Ly}^d
4[ms]	4[ms]	4[ms]	4[ms]	4[ms]	4[ms]
d_R	d_L	$\Delta\theta_R$	$\Delta\theta_L$	\mathbf{a}_{IMU}	\mathbf{x}_{VO}
4[ms]	4[ms]	4[ms]	4[ms]	4[ms]	10[ms]

Values marked with * have latency as same as sampling period.

(A) オンライン学習制御系

この系は PID 制御器に反復フィードバックチューニング (IFT) [81][171] を利用した学習制御系を組み合わせたシステムであり、さらに図 6.6 の制御システムが稼働中（ロボット実機が行動中）の状態においてもオンラインで PID ゲインを自動的に更新するチューニング機能を備えている。オンライン更新時には、ロボットの前後方向（平行二輪電動スクーター相対での X 方向）の並進速度 v に対する PID ゲイン $K_p^v(n), K_I^v(n), K_D^v(n)$ (n : 更新番号) と、ロボットの Yaw まわり（平行二輪電動スクーター相対での Z まわり）の旋回角速度 ω に対する PID ゲイン $K_p^\omega(n), K_I^\omega(n), K_D^\omega(n)$ の値セットの他、I ゲイン $K_I^v(n)$ や $K_I^\omega(n)$ が増加した場合に補正機能を働かせる特殊な積分器を作動させる。

これらのブロックは可変の意味をもたせて図 6.6 内において斜め矢印を付加して表現しており、6.3 節で以下に示すそれぞれの機能について詳説する。

- ・ 図 6.6 -緑ブロック： Mini-batch-based closed-loop Learning Control (MLC)
- ・ 図 6.6 -赤ブロック： 前回値補正積分器

PID 制御器単体での役割は、並進速度 v 、旋回角速度 ω とそれぞれの目標値 v^d, ω^d との偏差を $e_v (= v^d - v)$ 、 $e_\omega (= \omega^d - \omega)$ 、ロボット足部の Pitch 成分の傾斜角を微修正するための操作量を u_v, u_ω として、(D) ロボットプラントに与えている。コントローラの伝達関数 $C_*(s)$ ($* = v, \omega$) は以下の通りである。なお、偏差 e_v, e_ω は学習制御則においても使用する。

$$u_* = C_*(s)e_* = \left(K_P^*(n) + \frac{K_I^*(n)}{s} + K_D^*(n)s \right) e_*, \quad * = v, \omega \quad (6.2)$$

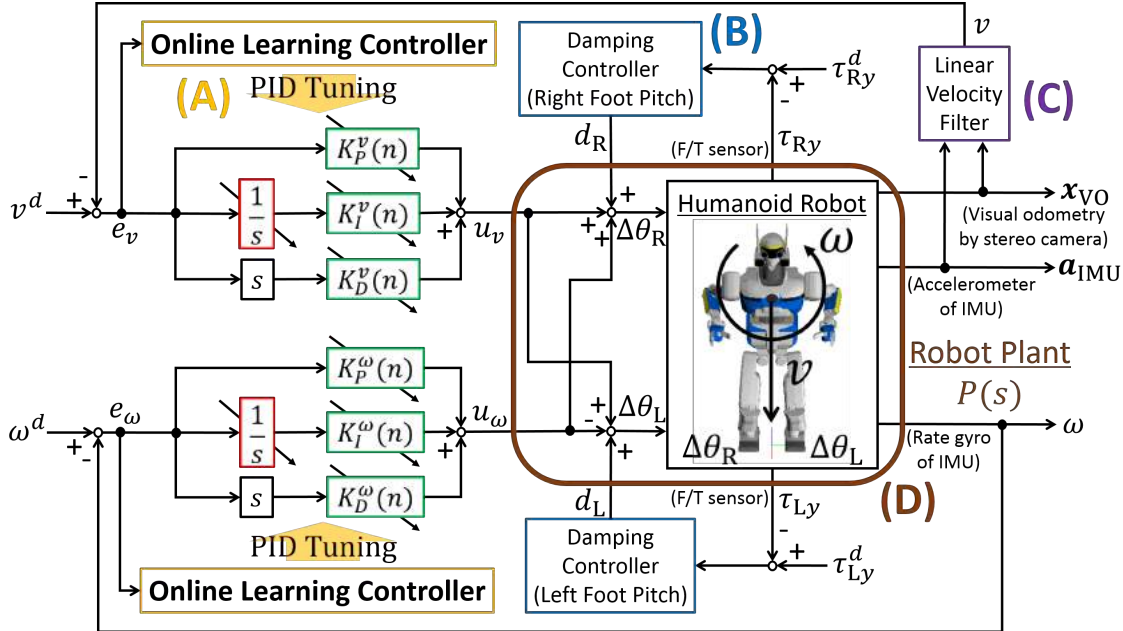


図 6.6: Speed Governing Behavior Control System on Parallel Two-Wheeled Electric Scooter
 This system is composed of (A) Online Learning Controller, (B) Foot Torque Damping Controller, (C) Linear Velocity Filter, and (D) Robot Plant.

(B) 足部トルクダンピング制御系

この系は 5.2 節で詳説したシステムであり，図 6.6 のシステムにおいては平行二輪電動スクーターの操作に関する Pitch 成分のみについて書き表している．足部トルクダンピング制御系は，調速行動制御に直接寄与しないバランス安定化（行動レベル 2）としての立ち位置であり，図 6.6 のフィードバック制御系においては式 5.3 の制御入力 δ を，左右の足部の Pitch 傾斜角に対する修正量に関して外乱 d_R, d_L として取り扱う．なお，足部トルクダンピング制御系のダンピングゲイン D は 5.3 節で述べた SLC により事前に学習されるパラメータである．

(C) 並進速度フィルタ

調速行動制御における制御量（状態量）は並進速度 v と旋回角速度 ω である．ヒューマノイドロボットの多くは姿勢推定などの用途で体内に IMU を備えていることが多く，IMU のうちの 3 軸ジャイロセンサにより微分演算や積分演算を必要とすることなく 3 次元の角速

度が得られる．一方で，並進速度はIMUのみから取得することが難しく，IMUのうちの加速度センサから計測できる並進加速度を時間積分するとノイズ誤差によるドリフトが発生することから，ドリフト補正のために別のセンサを加えて相補フィルタ [180][181] やカルマンフィルタ [182] を適用した状態推定を行う方法が知られている．

本研究ではYawまわりの旋回角速度 ω をIMUのジャイロセンサから純粹に取得し，X方向の並進速度 v をIMUの3軸加速度センサの計測値 \mathbf{a}_{IMU} とビジュアルオドメトリを利用したヒューマノイドの自己位置推定による計測値 \mathbf{x}_{VO} を統合するフィルタ処理により状態量を取得する．本研究ではこのフィルタを並進速度フィルタと呼ぶ．

ヒューマノイドロボットの自己位置推定では，Stereolabs [183] が提供するステレオカメラ“ZED” (図6.7) をロボット頭部に取り付け，3次元位置 X, Y, Z のビジュアルオドメトリ \mathbf{x}_{VO} を10[ms]周期でROS (Robot Operating System) [140] による実時間通信として計測する．並進速度フィルタは図6.8のような構成となっており， \mathbf{a}_{IMU} と \mathbf{x}_{VO} を入力して v を出力する． \mathbf{a}_{IMU} から重力加速度のオフセット \mathbf{g}_0 を除去した $\mathbf{a}_{\text{IMU}} - \mathbf{g}_0$ と \mathbf{x}_{VO} の時間微分 $d\mathbf{x}_{\text{VO}}/dt$ それぞれをロボットのローカル座標系X成分で表した a_x と v_x を求めた上で， a_x, v_x 各々を相補フィルタ (フィルタの時定数を T_f とする) により統合することで状態量 v を推定する (式6.3) ．

$$v = \frac{T_f s}{1 + T_f s} a_x + \frac{1}{1 + T_f s} v_x \quad (6.3)$$

相補フィルタは以下の2つのフィルタの信号処理を重ね合わせることで，加速度の時間積分によるドリフトや位置の時間微分によるノイズ増幅を補正し合う効果を得る．

- ハイパスフィルタ (HPF)： 加速度 a_x の高周波領域のみを通過させる
- ローパスフィルタ (LPF)： 位置の時間微分から得られる v_x の低周波領域のみを通過させる

(D) ロボットプラント

図6.6中央の褐色の囲み部分をロボットプラントと定め，ロボットプラントの伝達関数行列 $\mathbf{P}(s)$ を2入力 $[u_v \ u_\omega]^T$ ，2出力 $[v \ \omega]^T$ として以下に定義する．

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \mathbf{P}(s) \begin{bmatrix} u_v \\ u_\omega \end{bmatrix} = \begin{bmatrix} P_v(s) & P_{v\omega}(s) \\ P_{\omega v}(s) & P_\omega(s) \end{bmatrix} \begin{bmatrix} u_v \\ u_\omega \end{bmatrix} \quad (6.4)$$



図 6.7: Stereo Camera “ZED” [183]

ロボットプラント内でヒューマノイドに直接送る指令は，足部の Pitch 傾斜角 θ_R, θ_L の変位量 $\Delta\theta_R, \Delta\theta_L$ である． $\Delta\theta_R, \Delta\theta_L$ は，並進速度の制御入力 u_v と旋回角速度の制御入力 u_ω ，ならびに (B) 足部トルクダンピング制御系からの外乱 d_R, d_L により最終的にまとめて与えられる修正量として以下の式で記述できる．

$$\begin{bmatrix} \Delta\theta_R \\ \Delta\theta_L \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u_v \\ u_\omega \end{bmatrix} + \begin{bmatrix} d_R \\ d_L \end{bmatrix} \quad (6.5)$$

次にロボットに与えられる修正量 $[\Delta\theta_R \ \Delta\theta_L]^T$ とロボットから出力される状態量 $[v \ \omega]^T$ の関係について考える． $[\theta_R \ \theta_L]^T$ と $[v \ \omega]^T$ の関係は式 A.4 で与えられ，式 A.4 の両辺を時間微分することで以下を得る．

$$\begin{bmatrix} \dot{v} \\ \dot{\omega} \end{bmatrix} = r \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{W} & -\frac{1}{W} \end{bmatrix} \begin{bmatrix} K_R & 0 \\ 0 & K_L \end{bmatrix} \begin{bmatrix} \dot{\theta}_R \\ \dot{\theta}_L \end{bmatrix} \quad (6.6)$$

制御周期は $\Delta t = 4[\text{ms}]$ と十分小さく，計算機上では

$$\begin{bmatrix} \dot{\theta}_R \\ \dot{\theta}_L \end{bmatrix} \approx \begin{bmatrix} \frac{\Delta\theta_R}{\Delta t} \\ \frac{\Delta\theta_L}{\Delta t} \end{bmatrix} \quad (6.7)$$

と近似でき，また微分演算によるノイズ増幅をカットする用途でローパスフィルタをかけることで以下が得られる．

$$\begin{bmatrix} \dot{v} \\ \dot{\omega} \end{bmatrix} = \text{LPF}_{T_v, T_\omega} \left\{ \frac{r}{\Delta t} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{W} & -\frac{1}{W} \end{bmatrix} \begin{bmatrix} K_R & 0 \\ 0 & K_L \end{bmatrix} \begin{bmatrix} \Delta\theta_R \\ \Delta\theta_L \end{bmatrix} \right\} \quad (6.8)$$

$\text{LPF}_{T_v, T_\omega} \{ \}$ は 1 行目，2 行目それぞれを時定数 T_v, T_ω でローパスフィルタ演算する関数である．式 6.8 の両辺をラプラス変換して次式を得る．

$$\begin{bmatrix} sv \\ s\omega \end{bmatrix} = \frac{r}{\Delta t} \begin{bmatrix} \frac{1}{1+T_v s} & 0 \\ 0 & \frac{1}{1+T_\omega s} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{W} & -\frac{1}{W} \end{bmatrix} \begin{bmatrix} K_R & 0 \\ 0 & K_L \end{bmatrix} \begin{bmatrix} \Delta\theta_R \\ \Delta\theta_L \end{bmatrix} \quad (6.9)$$

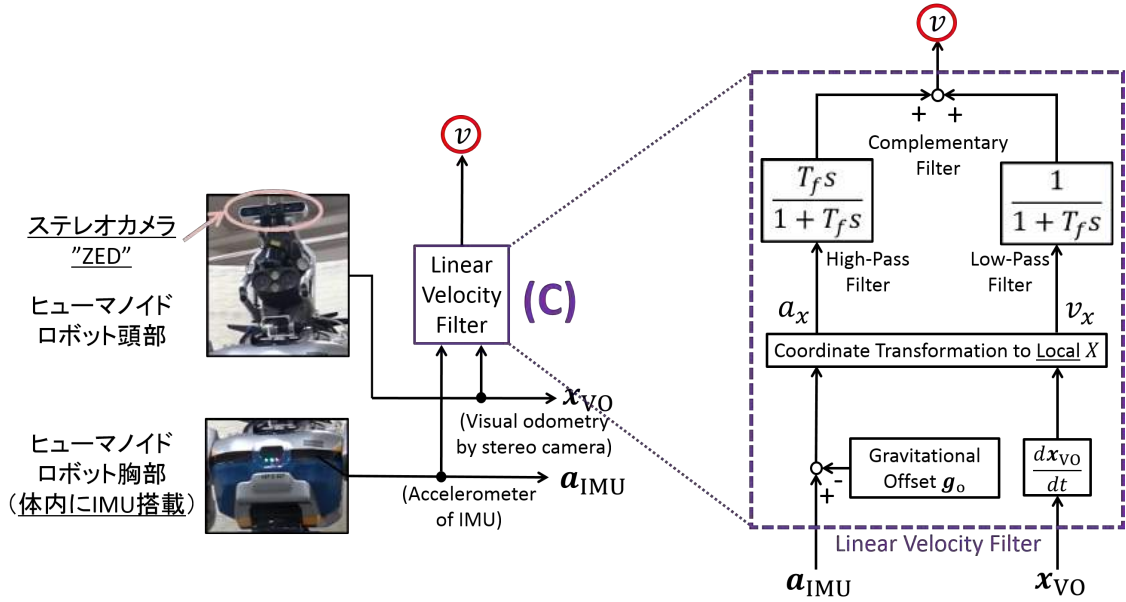


図 6.8: Linear Velocity Filter Configuration

式 6.5 の外乱項 $[d_R \ d_L]^T$ を除いて式 6.9 に代入して整理することで $[u_v \ u_\omega]^T$ と $[v \ \omega]^T$ の関係は以下で表される。

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} \frac{r}{\Delta t} \cdot \frac{K_R + K_L}{2} \cdot \frac{1}{s(1 + T_v s)} & \frac{r}{\Delta t} \cdot \frac{K_R - K_L}{2} \cdot \frac{1}{s(1 + T_v s)} \\ \frac{r}{\Delta t} \cdot \frac{K_R - K_L}{W} \cdot \frac{1}{s(1 + T_\omega s)} & \frac{r}{\Delta t} \cdot \frac{K_R + K_L}{W} \cdot \frac{1}{s(1 + T_\omega s)} \end{bmatrix} \begin{bmatrix} u_v \\ u_\omega \end{bmatrix} \quad (6.10)$$

図 A.6 の比例特性の結果より、

$$K_R - K_L \approx 0 \quad (6.11)$$

とみなせ、定数項を K_v および K_ω とおくことで式 6.4 で定義した伝達関数行列 $\mathbf{P}(s)$ は以下と求まる。

$$\mathbf{P}(s) = \begin{bmatrix} P_v(s) & P_{v\omega}(s) \\ P_{\omega v}(s) & P_\omega(s) \end{bmatrix} = \begin{bmatrix} \frac{K_v}{s(1 + T_v s)} & 0 \\ 0 & \frac{K_\omega}{s(1 + T_\omega s)} \end{bmatrix} \quad (6.12)$$

$\mathbf{P}(s)$ は対角行列であり、 v, ω 共通で

$$P_*(s) = \frac{K_*}{s(1 + T_* s)}, \quad * = v, \omega \quad (6.13)$$

とプラントの伝達関数は表現できる。

式 6.12 の $P(s)$ を式 6.4 に代入して時間領域に直すと次式の関係となる.

$$\begin{bmatrix} \dot{v} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} K_v & 0 \\ 0 & K_\omega \end{bmatrix} \begin{bmatrix} \text{LPF}_{T_v}\{u_v\} \\ \text{LPF}_{T_\omega}\{u_\omega\} \end{bmatrix} \quad (6.14)$$

式 6.14 は入力 $[u_v \ u_\omega]^T$ と出力 $[\dot{v} \ \dot{\omega}]^T$ のモデルを表しており, モデルの出力データ $[\dot{v} \ \dot{\omega}]^T$ と実際の計測データ $[\dot{v}_{\text{act}} \ \dot{\omega}_{\text{act}}]^T$ が合うようにモデルパラメータ $T_v, T_\omega, K_v, K_\omega$ を決定する. 入力データ $[u_v \ u_\omega]^T$ と計測データ $[\dot{v}_{\text{act}} \ \dot{\omega}_{\text{act}}]^T$ の位相が揃うようにローパスフィルタ (LPF) の一次遅れ系の時定数 T_v, T_ω を決めた後, モデル化誤差の二乗和

$$\sum_t (\dot{v}_{\text{act}} - \dot{v})^2 \quad (6.15)$$

$$\sum_t (\dot{\omega}_{\text{act}} - \dot{\omega})^2 \quad (6.16)$$

が最小となるように K_v, K_ω を最小二乗法により同定する.

同定結果を表 6.2 および図 6.9 に示す. 図 6.9 において, 赤線が式 6.14 によるモデルデータ $\dot{v}, \dot{\omega}$, 緑線が実際に計測されたデータ $\dot{v}_{\text{act}}, \dot{\omega}_{\text{act}}$ を表す. $\dot{v}_{\text{act}}, \dot{\omega}_{\text{act}}$ は計測値 v, ω の信号の時間微分演算により高周波ノイズが含まれる. (a) の 1[s], 24[s] 付近, (b) の 1[s], 3.5[s], 6.5[s], 9.5[s], 12.5[s] 付近など定常状態において残差 (オフセット) が残っているものの, 変化が観察できる部分については概ねモデルデータと実際の計測データが揃っており同定に成功している. 位相を揃える際の時定数 T_v, T_ω の僅かなズレや, $\dot{v}_{\text{act}}, \dot{\omega}_{\text{act}}$ に含まれる高周波ノイズ等が残差の残留に影響していると考察される. 導出したプラントの伝達関数 $P_*(s) = \frac{K_*}{s(1+T_*s)}$ ($*$ = v, ω) ならびにパラメータ同定結果 (表 6.2) は 6.3 節の学習制御則において用いる.

表 6.2: Identified Values for Parameters T_*, K_*

T_v	T_ω	K_v	K_ω
1.1	0.6	5.88×10^2	1.39×10^4

6.3 PID ゲインのオンライン学習制御

本節では, 図 6.6 の调速行動制御系のシステム構成における (A) オンライン学習制御系の具体的な構成法を述べ, 主に反復フィードバックチューニング (IFT) を利用した PID ゲインの学習則として “Mini-batch-based closed-loop Learning Control (MLC)” の手法と, I ゲイ

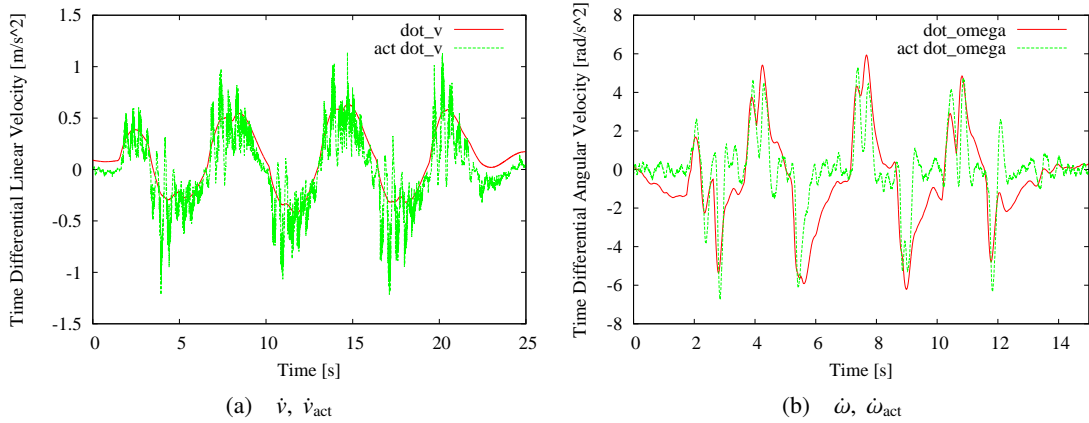


図 6.9: Model Identification Results of $[\dot{v} \ \dot{\omega}]^T$ (Red Solid Lines) Compared with the Actual Measured Data $[\dot{v}_{act} \ \dot{\omega}_{act}]^T$ (Green Dashed Lines)

ンをオンラインでチューニングする際に問題となる偏差積分値の蓄積の解消を目的とした“前回値補正積分器”を提案し，調速行動制御系に対するオンライン学習制御法について論じる．以降の議論では， $*$ = v, ω として 2 つの状態量を共通に取り扱うものとする．

6.3.1 Mini-batch-based closed-loop Learning Control (MLC)

5.3 節で提案した搭乗する環境への適応を目的とした即時的な学習制御法“SGD-based open-loop Learning Control (SLC)”に対して，搭乗型機器への乗り込み後などの搭乗した環境の運動への適応を目的とした学習制御法を“Mini-batch-based closed-loop Learning Control (MLC)”と定義する（図 6.10 に示すアプローチ）．SLC と MLC の主な違いを表 6.3 に示す．

表 6.3: SLC と MLC の学習制御法の違い

	SLC	MLC
学習データの処理形式	online learning (SGD)	mini-batch learning
制御系の取り扱い	open-loop	closed-loop

瞬間的な乗り込み動作では SLC により迅速なゲイン学習を実行しているのに対して，乗り込み後の調速行動に適用される MLC では即時性よりも厳密性を考慮して制御系を closed-loop として扱っている．また，学習データの処理方法に関しては，SLC は SGD に基づいたオンライン学習形式で実行されており，図 5.4 で示したように実制御時間毎に更新を行うのに対して，MLC はある程度の学習データの蓄積をとり，平均的な情報を利用している．

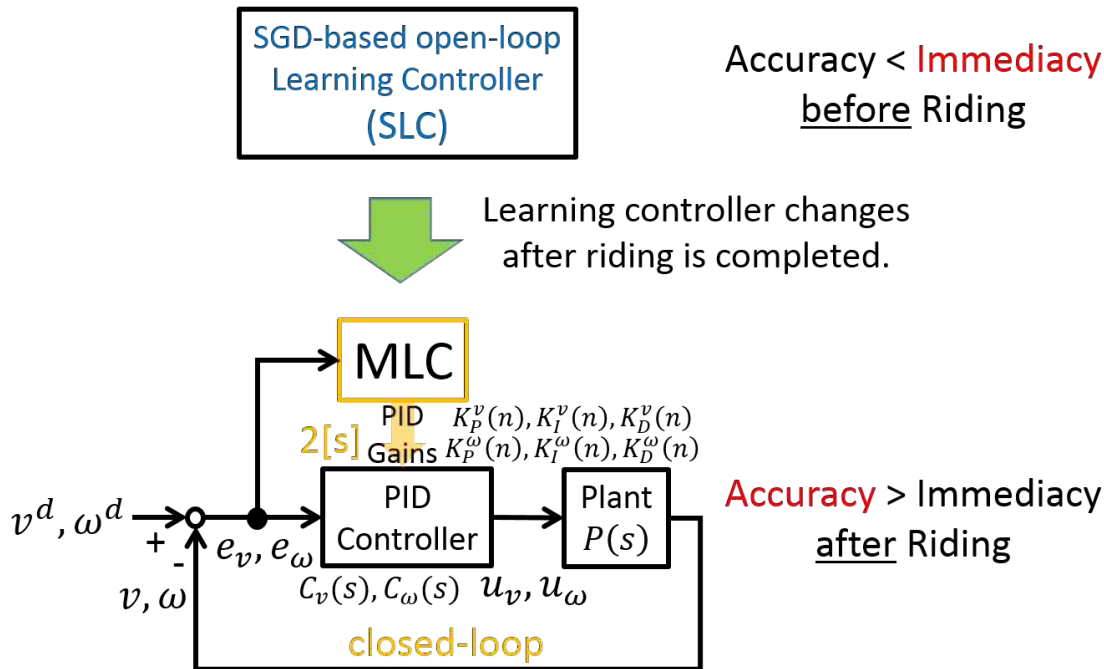


図 6.10: MLC Approach

SLC と MLC どちらの処理形式もオンラインでの学習処理方法であるが、後者の MLC は一定期間（本研究においては 2[s] 毎）のデータ蓄積による待機のフェーズが存在することから制御周期毎の更新ではなく、ミニバッチサイズ毎の更新となる。このことから、MLC の名称定義には“Mini-batch-based”という単語が含まれている。

以下で MLC の具体的なゲイン学習法のアルゴリズムについて述べる。

6.3.2 MLC によるゲイン学習法

PID コントローラをはじめとするフィードバック制御器の最適ゲインを探索する手法として、反復フィードバックチューニング (IFT) [81][171] を用いた反復的な学習制御法がよく知られている¹⁾。本研究で提案した調速行動制御系においても PID コントローラが活用されており、IFT の適用は有効であるといえる。MLC によるゲイン学習の基本則として IFT を採用することで、計算機上（離散時間系）においてオンラインに PID ゲインが学習可能な形

¹⁾IFT の詳細なアルゴリズムや安定性・ロバスト性の解析については、文献 [81][171] に譲る。

式として表す.

評価関数の設定

IFTの基本となるアルゴリズムは、最適化したいパラメータベクトル ρ の評価関数 $J(\rho)$ を最小化する局所解 $\rho = \rho_{\text{opt}}$ を勾配法（最急降下法）探索により求めることであり、パラメータに対する評価関数の勾配 $\nabla J(\rho)$ が計算できればよい. IFTなどの学習制御では、評価関数として制御量 $y(t)$ と目標値 $y^d(t)$ の偏差 $e_*(t)$ の時間 t に対する二乗和（積分値）が与えられることが多く、本論文でも以下のように評価関数を設ける.

$$J(\rho_n) = \frac{1}{T_L} \int_t^{t+T_L} (y^d(t) - y(t))^2 dt = \frac{1}{T_L} \int_t^{t+T_L} e_*(t)^2 dt, \quad \rho_n = [K_P^*(n) \ K_I^*(n) \ K_D^*(n)]^T \quad (6.17)$$

ここでの $y(t)$ と $y^d(t)$ は v, ω と v^d, ω^d の状態量に相当する. $\rho_n \in \mathbb{R}^3$ はゲイン更新 n 回目でのPIDゲインを並べたパラメータベクトルであり、 T_L は評価関数計算のサンプリング周期で1回あたりの勾配計算に要する学習時間（学習制御の更新周期）に相当するミニバッチ処理の設定パラメータである. T_L が小さ過ぎると1更新あたりの学習が迅速に終了する一方で、少ないサンプリングで評価関数計算することになり、また T_L が大き過ぎても学習に時間がかかってしまう. 実験ではこれらのトレードオフを考慮して $T_L = 2[\text{s}]$ （ミニバッチサイズ： $2/0.004 = 500$ ）に設定している.

学習勾配の導出

式6.17の評価関数 $J(\rho_n)$ をパラメータベクトル ρ_n で偏微分することで学習勾配 $\nabla J(\rho_n)$ は以下と求まる.

$$\nabla J(\rho_n) = \left[\frac{\partial J(\rho_n)}{\partial \rho_n} \right]^T = -\frac{2}{T_L} \int_t^{t+T_L} \left[\frac{\partial y(t)}{\partial \rho_n} \right]^T e_*(t) dt \quad (6.18)$$

ここで、 $[\partial y(t)/\partial \rho_n]^T$ のラプラス変換 $[\partial y(s)/\partial \rho_n]^T$ はコントローラとプラントの伝達関数 $C_*(s), P_*(s)$ を用いて次式で表すことができる [81][171].

$$\left[\frac{\partial y(s)}{\partial \rho_n} \right]^T = \frac{P_*(s)}{1 + P_*(s)C_*(s)} \left[\frac{\partial C_*(s)}{\partial \rho_n} \right]^T e_*(s) \quad (6.19)$$

$C_*(s) = K_P^*(n) + \frac{K_I^*(n)}{s} + K_D^*(n)s$ および $P_*(s) = \frac{K_*}{s(1+T_*s)}$ を代入して $[\partial C_*(s)/\partial \rho_n]^T = [1 \ \frac{1}{s} \ s]^T$ となることから式6.19を s について整理すると以下となる.

$$\left[\frac{\partial y(s)}{\partial \rho_n} \right]^T = \frac{K_*}{T_*s^3 + (1 + K_*K_D^*(n))s^2 + K_*K_P^*(n)s + K_*K_I^*(n)} [s \ 1 \ s^2]^T e_*(s) \quad (6.20)$$

$[\partial y(t)/\partial \rho_n]^T = \mathbf{Y}(t) \in \mathbb{R}^3$ において式 6.20 を時間領域で表した微分方程式として記述する.

$$T_* \mathbf{Y}^{(3)}(t) + (1 + K_* K_D^*(n)) \mathbf{Y}^{(2)}(t) + K_* K_P^*(n) \mathbf{Y}^{(1)}(t) + K_* K_I^*(n) \mathbf{Y}(t) = K_* [\dot{e}_*(t) \ e_*(t) \ \ddot{e}_*(t)]^T \quad (6.21)$$

ここで, $\mathbf{Y}^{(m)}(t)$ は $\mathbf{Y}(t)$ の m 階時間微分を表す.

$$\mathbf{Y}^{(1)}(t) = \frac{d\mathbf{Y}(t)}{dt} = \frac{1}{\Delta t} (\mathbf{Y}(t) - \mathbf{Y}(t - \Delta t)) \quad (6.22)$$

$$\mathbf{Y}^{(2)}(t) = \frac{d\mathbf{Y}^{(1)}(t)}{dt} = \frac{1}{\Delta t^2} (\mathbf{Y}(t) - 2\mathbf{Y}(t - \Delta t) + \mathbf{Y}(t - 2\Delta t)) \quad (6.23)$$

$$\mathbf{Y}^{(3)}(t) = \frac{d\mathbf{Y}^{(2)}(t)}{dt} = \frac{1}{\Delta t^3} (\mathbf{Y}(t) - 3\mathbf{Y}(t - \Delta t) + 3\mathbf{Y}(t - 2\Delta t) - \mathbf{Y}(t - 3\Delta t)) \quad (6.24)$$

のそれぞれを微分方程式 6.21 へ代入して差分方程式に変換することで $\mathbf{Y}(t)$ に関する離散時間の更新式が得られる.

$$\mathbf{Y}(t) = \frac{1}{C_0} (C_1 \mathbf{Y}(t - \Delta t) + C_2 \mathbf{Y}(t - 2\Delta t) + C_3 \mathbf{Y}(t - 3\Delta t) + K_* [\dot{e}_*(t) \ e_*(t) \ \ddot{e}_*(t)]^T) \quad (6.25)$$

$$C_0 = \frac{T_*}{\Delta t^3} + \frac{1 + K_* K_D^*(n)}{\Delta t^2} + \frac{K_* K_P^*(n)}{\Delta t} + K_* K_I^*(n) \quad (6.26)$$

$$C_1 = \frac{3T_*}{\Delta t^3} + \frac{2(1 + K_* K_D^*(n))}{\Delta t^2} + \frac{K_* K_P^*(n)}{\Delta t} \quad (6.27)$$

$$C_2 = -\frac{3T_*}{\Delta t^3} - \frac{1 + K_* K_D^*(n)}{\Delta t^2} \quad (6.28)$$

$$C_3 = \frac{T_*}{\Delta t^3} \quad (6.29)$$

以上より, 表 6.2 の T_* , K_* の同定結果を用いて更新式 6.25 より逐次的に $\mathbf{Y}(t) = [\partial y(t)/\partial \rho_n]^T$ を演算し, 式 6.18 に代入して積分することで学習勾配 $\nabla J(\rho_n)$ を計算機上でオンラインに得ることが可能となる.

PID ゲイン列 ρ_n の更新則

学習勾配 $\nabla J(\rho_n)$ を用いて PID ゲインのパラメータベクトル ρ_n を次の式により更新する.

$$\rho_{n+1} := \rho_n - \text{diag}(\eta_P, \eta_I, \eta_D) \nabla J(\rho_n) \quad (6.30)$$

ここで, $\text{diag}(\eta_P, \eta_I, \eta_D) \in \mathbb{R}^{3 \times 3}$ は勾配法 (最急降下法) における各 PID ゲインの学習率 η_P, η_I, η_D からなる対角行列である.

最後に, 学習で得られた ρ_{n+1} が PID ゲインとして非負値をとっているかを更新毎にチェックし, 負値の結果ならば 0 を割り当てて再学習を行う.

オンライン探索指標に基づくゲイン学習のアプローチ

上記に述べた MLC によるゲイン学習の処理についてオンライン探索指標の側面から考える。5.3.2 項で議論したアプローチと同様に、評価関数 $J(\mathbf{x}(t; n))$ の引数である探索変数 $\mathbf{x}(t; n)$ について調べる。定義した評価関数 $J(\boldsymbol{\rho}_n)$ (式 6.17) を式 2.10 で示した二次形式の表現に置き換えると、探索変数 $\mathbf{x}(t; n)$ 、探索目標 \mathbf{x}_d 、および重み行列 W はそれぞれ以下の対応となる。

$$\mathbf{x}(t; n) = y(t), \quad \mathbf{x}_d = y^d(t), \quad W = \frac{1}{T_L} \quad (6.31)$$

式 6.17 で定義した評価関数 $J(\boldsymbol{\rho}_n)$ は $\mathbf{x}(t; n) = \mathbf{x}_d$ で最小となり、式 2.9 の条件式を満たしていることから、 $y(t) = v, \omega$ 、 $y^d(t) = v^d, \omega^d$ 、式 6.30 よりオンライン探索指標に基づくゲイン学習の一般化更新則 (式 2.15) との対応は以下となる。

表 6.4: MLC におけるオンライン探索指標およびゲイン学習のパラメータの対応

探索変数	探索目標
v, ω	v^d, ω^d
更新されるゲイン	一般化更新則の重み A
$\boldsymbol{\rho}_n$	$\text{diag}(\eta_P, \eta_I, \eta_D)$

$\eta_P > 0, \eta_I > 0, \eta_D > 0$ であることから、 $A = \text{diag}(\eta_P, \eta_I, \eta_D)$ は正定値対称行列である。

終了条件

以上の MLC の学習制御の終了条件は以下の式が成立した時点とする。

$$\|\nabla J(\boldsymbol{\rho}_n)\|_2 < \varepsilon \quad \wedge \quad J(\boldsymbol{\rho}_n) < \sigma \quad (t_0 \leq t \leq t_0 + kT_L) \quad (6.32)$$

前者の条件 $\|\nabla J(\boldsymbol{\rho}_n)\|_2 < \varepsilon$ は勾配の L^2 ノルムが 0 近傍となる極小値の収束条件であり、実際の計算機環境では極めて小さい正値をとる ε で抑えられた条件とする。後者の条件 $J(\boldsymbol{\rho}_n) < \sigma$ は評価関数 $J(\boldsymbol{\rho}_n)$ が 0 近傍となる、制御系の偏差が 0 に近くなった場合の収束条件であり、同様に小さな正値 σ で抑えられた条件としている。これらの 2 つの条件の同時成立に加えて、時間に対しての定常的な条件として $t_0 \leq t \leq t_0 + kT_L$ の連続した学習 k 回分の期間について上記条件を満たすとき学習制御を終了する。実験では $k = 3$ と設定している。

6.3.3 前回値補正積分器

図 6.6 の赤枠で囲まれた積分器ブロックを前回値補正積分器と呼び、MLC で学習されたゲイン結果を PID 制御器に反映（更新）する際に、I ゲインの変更が制御入力の発散に影響するのを防ぐ働きをもつ。前回値補正積分器は、計算機上で偏差を積算する際、デジタルの積分器がもつ切替スイッチにより前回（前時刻）の積分結果に修正を加えて、I ゲインのチューニングに適応する能力がある。

ブロック線図における一般的な積分器の記法は図 6.11 -左上のようなブロックで表すが、本研究では偏差 e_* と、その積分値 E_* の前回値（現時刻の前までの積分値）の 2 入力のポートで表した図 6.11 -右上に示す等価な積分器ブロックとして定義する。なお、 z 変換のブロックを用いて表した等価な離散時間積分器は図 6.11 -下部で表される。これらの記法を基に、前回値補正積分器を離散時間系のブロック線図で表した図が図 6.12 である。以降の議論では、計算機上で実装可能な形式として離散時間系で記述するが、I ゲインの逐次的な更新 $K_I^*(n) \rightarrow K_I^*(n+1)$ の形式に倣い、 z 変換表現ではなく時間変化 $t \rightarrow t + \Delta t$ の表現で扱う。

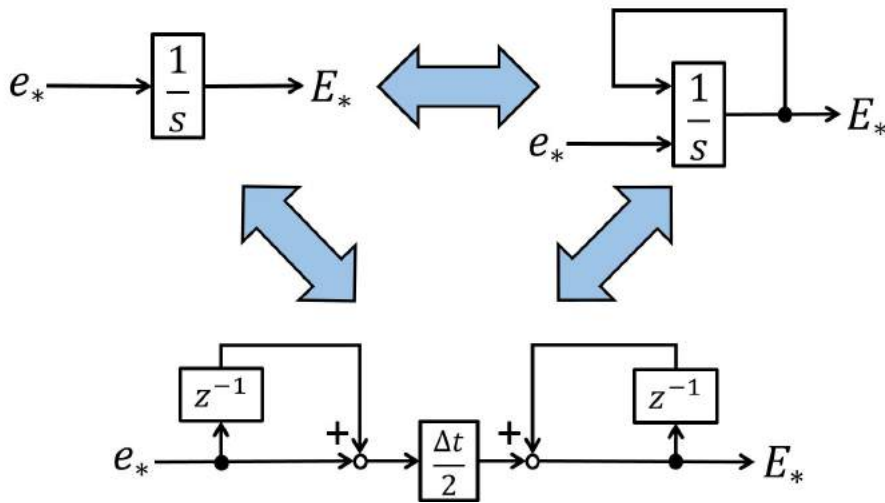


図 6.11: Equivalent Representation about the General Integrator

図 6.12 で示した前回値補正積分器の切替スイッチが SW0 の場合は、図 6.11 のように通常の積分器として振る舞う。このときの時刻 t における積分値 $E_*(t)$ と I 動作の制御入力 $u_I^*(t)$

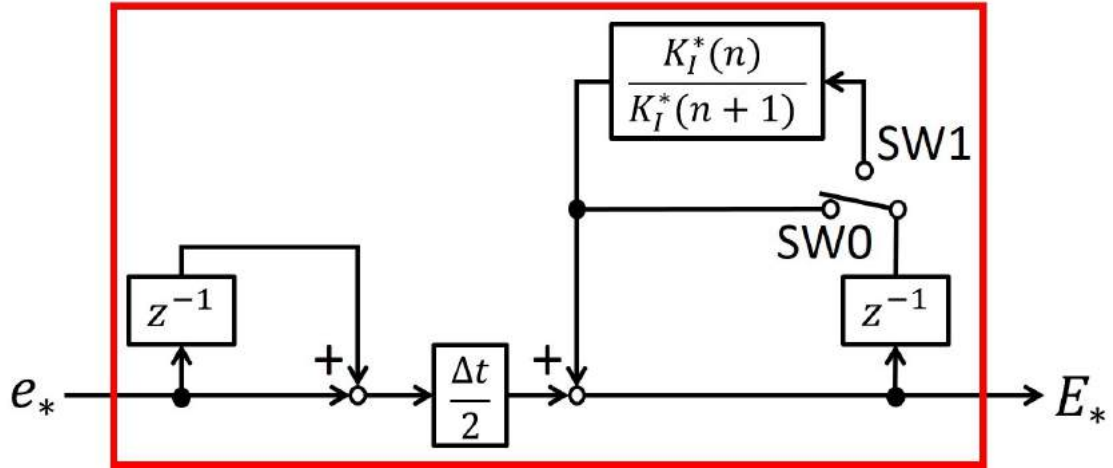


図 6.12: Discrete-Time Block Diagram of Proposed Integrator

は次式で演算される.

$$E_*(t) = E_*(t - \Delta t) + \frac{\Delta t}{2} (e_*(t) + e_*(t - \Delta t)) \quad (6.33)$$

$$u_I^*(t) = K_I^*(n)E_*(t) \quad (6.34)$$

時間変化 $t \rightarrow t + \Delta t$ に対してゲイン更新がない ($K_I^*(n) \rightarrow K_I^*(n)$) とき, $u_I^*(t + \Delta t)$ は $u_I^*(t)$ を用いて以下で表せる.

$$u_I^*(t + \Delta t) = K_I^*(n)E_*(t + \Delta t) = u_I^*(t) + \frac{\Delta t}{2} (e_*(t + \Delta t) + e_*(t)) K_I^*(n) \quad (6.35)$$

次に, 時間変化 $t \rightarrow t + \Delta t$ と同時に I ゲインが $K_I^*(n) \rightarrow K_I^*(n+1)$ に更新されるとき, $u_I^*(t + \Delta t)$ は以下となる.

$$u_I^*(t + \Delta t) = K_I^*(n+1)E_*(t + \Delta t) = \frac{K_I^*(n+1)}{K_I^*(n)} u_I^*(t) + \frac{\Delta t}{2} (e_*(t + \Delta t) + e_*(t)) K_I^*(n+1) \quad (6.36)$$

式 6.36 は式 6.35 と比較して, 時刻 t での制御入力 $u_I^*(t)$ が I ゲイン更新に応じて $\frac{K_I^*(n+1)}{K_I^*(n)}$ 倍されている. $\frac{K_I^*(n+1)}{K_I^*(n)} > 1$ であれば制御入力を増幅していることとなり, 特に $K_I^*(n) = 0$, $K_I^*(n+1) > 0$ のように変化が起こる場合は制御入力が発散してしまう. $K_I^*(n) = 0$, $K_I^*(n+1) > 0$ の例では, $K_I^*(n) = 0$ で修正しきれていない偏差の積分値が $K_I^*(n+1) > 0$ に更新された瞬間に大きな I 動作として働くため, 過大なオーバーシュートによる暴走の危険性がある. そ

の対策として、制御入力に飽和制約を設定して頭打ちさせた状態でのアンチwindアップ [172][173][174][175][176] の手法がよく知られているが、本研究ではPIDゲインの変化に応じた飽和のない正規の制御入力を与えることで挙動（状態量）を観測し学習させる目的で飽和制約を設けずに上記の問題を解決する。

本研究では、Iゲイン変化 $K_I^*(n) \rightarrow K_I^*(n+1)$ に伴う式 6.36 の $u_I^*(t)$ の増幅率 $\frac{K_I^*(n+1)}{K_I^*(n)}$ が制御入力の増大・発散に寄与するとして、増幅条件 $\frac{K_I^*(n+1)}{K_I^*(n)} > 1$ の場合の更新時に前回値補正積分器の切替スイッチを SW1 にして積分演算を以下のように切り替えることで対処する。

$$E_*(t) = \frac{K_I^*(n)}{K_I^*(n+1)} E_*(t-\Delta t) + \frac{\Delta t}{2} (e_*(t) + e_*(t-\Delta t)) \quad (6.37)$$

式 6.37 より時間変化 $t \rightarrow t + \Delta t$ 、Iゲイン変化 $K_I^*(n) \rightarrow K_I^*(n+1)$ でのI動作の制御入力 $u_I^*(t + \Delta t)$ は以下となる。

$$u_I^*(t + \Delta t) = K_I^*(n+1) E_*(t + \Delta t) = u_I^*(t) + \frac{\Delta t}{2} (e_*(t + \Delta t) + e_*(t)) K_I^*(n+1) \quad (6.38)$$

式 6.38 は式 6.35 の $K_I^*(n)$ を更新後の $K_I^*(n+1)$ に置き換えた形式であり、ゲイン更新に伴う制御入力 $u_I^*(t)$ の増幅を打ち消すことができる。増幅条件 $\frac{K_I^*(n+1)}{K_I^*(n)} > 1$ の場合は $K_I^*(n+1) > 0$ であることから、式 6.37 における $\frac{K_I^*(n)}{K_I^*(n+1)}$ は分母が0ではないことが保証されている。また、Iゲイン減少（または不変） $\frac{K_I^*(n+1)}{K_I^*(n)} \leq 1$ の場合は、Iゲインに依らない通常の積分器（式 6.33）として演算が実行されるため、減少の末に $K_I^*(n+1) = 0$ となる場合に対しても対処可能である。

以上より、増幅条件 $\frac{K_I^*(n+1)}{K_I^*(n)} > 1$ 、すなわち $K_I^*(n+1) > K_I^*(n)$ でゲインが更新される瞬間に前回値補正積分器を SW1 として（式 6.39）、式 6.37 の演算より積分量の前回値 $E_*(t - \Delta t)$ を $\frac{K_I^*(n)}{K_I^*(n+1)}$ 倍 ($0 \leq \frac{K_I^*(n)}{K_I^*(n+1)} < 1$) して補正することでI動作の発散を抑える。

$$SW = \begin{cases} 1 \text{ (eq.(6.37))}, & K_I^*(n+1) > K_I^*(n) \\ 0 \text{ (eq.(6.33))}, & K_I^*(n+1) \leq K_I^*(n) \quad \vee \text{ after switching} \end{cases} \quad (6.39)$$

ここで、SW1 へは $K_I^*(n+1) > K_I^*(n)$ の場合のゲイン更新時の一瞬にのみ切り替えを実行し、ゲイン更新（スイッチング）の完了後は次の更新時まで常に SW0（通常の積分器，式 6.33）であることに留意する。

前回値補正積分器の特徴として、飽和制約条件のリミットや目標値の修正などの設定が不要であり、更新前後のIゲイン $K_I^*(n), K_I^*(n+1)$ のみを用いて補正できる点が挙げられる。従来の対処法であるアンチwindアップ（Anti-windup） [172][173][174][175][176]、およ

び目標値整形・リファレンスガバナ (Reference shaping/governor) [177][178][179] と、本研究で提案した前回値補正積分器 (Proposed integrator) の違いを表 6.5 にまとめる。アンチウィンドアップ (Anti-windup) [172][173][174][175][176] の対処法では、コントローラの制御入力等に設けた飽和制約の閾値の決定や飽和量まわりでのゲイン設定が必要となる。目標値整形やリファレンスガバナ (Reference shaping/governor) [177][178][179] は、システムに存在する拘束や飽和制約を破ることなく目標値に整形を加える最大出力許容集合の構成 [184][185][186][187] や、最大 CPI (Constraint Positively Invariant) 集合 [184][188] の考えを用いることで ZMP に関する力学的拘束を満足させた状態フィードバックゲインの選択・スイッチングによる安定化制御 [169][189] などが提案されており、これらはシステムにおける拘束条件や与える飽和制約が既知であることが前提となる。これらの方策と比較して、前回値補正積分器は I ゲインの変更時にのみ有効な限定的な手法ではあるが、飽和制約の閾値設定や拘束条件の記述、目標値整形等を要さない偏差の補正処理が可能なシンプルな手法となっている。

さらに、前回値補正積分器は式 6.37 や式 6.38 で示したように $E_*(t - \Delta t) \rightarrow E_*(t)$, $u_1^*(t) \rightarrow u_1^*(t + \Delta t)$ に関する一般化更新則としての側面を持ち合わせており、軌道修正やゲイン学習の一部として考えることもできる。

表 6.5: Comparison of the Approach for Integral Action

	Anti-windup	Reference shaping/governor	Proposed integrator
Saturation limit	Necessary	Unnecessary	Unnecessary
Reference modification	Unnecessary	Necessary	Unnecessary
Execution timing	Control input is saturated	Judged from maximal sets	Integral gain increases
Required parameters	<ul style="list-style-type: none"> • Saturation thresholds • Saturation feedback gains 	<ul style="list-style-type: none"> • Constraint conditions • Finite numbers of iteration 	<ul style="list-style-type: none"> • Two integral gains before/after update

6.3.4 MLC および前回値補正積分器に基づくオンライン学習制御系

本節でこれまでに述べた MLC の学習制御処理と、I ゲインのチューニングをサポートする前回値補正積分器を統合したオンライン学習制御系の詳細なシステム構成を図 6.13 に示す。図 6.13 は、図 6.6 で示した调速行動制御系の全体構成の内の (A) の要素システム内部を展開

した構成図であり、 v と ω に対する閉ループ制御系を共通に表したものである。図6.13の緑枠のブロックがMLCに基づく処理系であり、赤枠のブロックが前回値補正積分器に基づく処理系である。2つの処理系を統合した全体のシステムにおけるPIDゲインの学習プロセスは以下の通りである。

1. 探索変数 v, ω と探索目標 v^d, ω^d の偏差 e_v, e_ω を学習データとして、 $T_L = 2[s]$ （データサイズ： $2/0.004 = 500$ ）分のミニバッチデータから学習勾配を計算。
2. 導出した学習勾配を基に、新規（更新後）のPIDゲイン
 $K_P^v(n+1), K_I^v(n+1), K_D^v(n+1), K_P^\omega(n+1), K_I^\omega(n+1), K_D^\omega(n+1)$
 の全6パラメータの値を決定。
3. 学習結果のPIDゲインの符号に対する例外処理/割り当て/再学習。
4. 前回値補正積分器によるIゲイン更新に対するスイッチング処理。
5. 実行中の制御系に対するPIDゲインのアップデート（数値代入/値セット）。
6. ミニバッチの学習時間 $T_L = 2[s]$ 周期で手順1～手順5をオンラインに実行。

6.4 平行二輪電動スクーター上での調速行動

本節では、6.2節で確立した調速行動制御系および6.3節で提案した調速行動制御系におけるPIDゲインのオンライン学習制御法を、等身大ヒューマノイドロボットHRP2-JSK [110]実機に適用することで、図6.14に示すような平行二輪電動スクーター上でのヒューマノイドの調速行動を実現する。

6.4.1 PIDゲインをオンラインで学習する調速行動実験

等身大ヒューマノイドが平行二輪電動スクーターに搭乗した環境での運動に適応しながら、実装された調速行動制御系のPIDゲインを初期値0より学習・獲得する調速行動実験を行った。調速行動の実験結果を図6.15，図6.16に示す。図6.15，図6.16はそれぞれ以下の実験条件のもと行った。

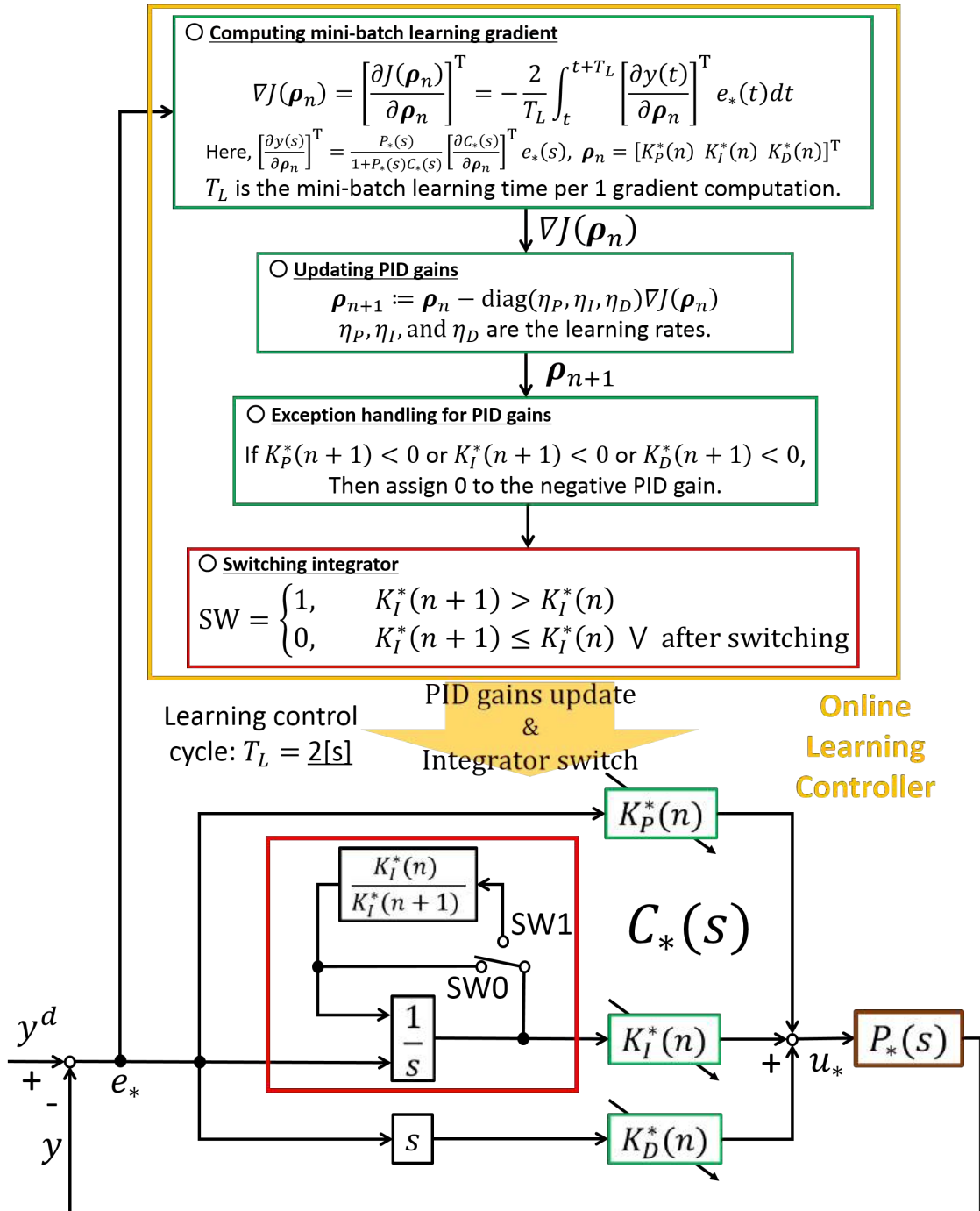


図 6.13: Online Learning Control System Based on MLC and Proposed Integrator



図 6.14: Speed Governing Behavior of Life-Sized Humanoid Robot on Parallel Two-Wheeled Electric Scooter

図 6.15：介護用リフトを用いて人間が手動で等身大ヒューマノイドを平行二輪電動スクーター上に乗せてから支えを外して学習を開始。

図 6.16：平坦な床面上から等身大ヒューマノイドが乗り込み動作を生成して自らが搭乗した後に学習を開始。

図 6.15 におけるオンライン学習制御の実行開始は、サポートにより搭乗した状態から人間が Joystick コントローラのスタートボタンを押下したタイミングとし、図 6.16 におけるオンライン学習制御の実行開始は、ヒューマノイドが床面上から平行二輪電動スクーター上へと乗り込み動作を経て両足がボード上に着地したタイミングとする。図 6.16 での学習開始となる着地状態の判定は、図 5.5 で示したような足裏反力の閾値判定に基づくものとする。

図 6.15，図 6.16 はともに学習の開始までは行動レベル 2 であり、調速行動制御系を OFF にした状態から開始する。図 6.16 における乗り込み動作生成では、バランス安定化制御 [55] に対して第 5 章で獲得した足部トルクダンピング制御系まわりのゲイン/パラメータを事前に適用した状態から実行する。

学習中は並進速度・旋回角速度が十分に観測できるよう、ロボットが十分に動き回れる程度の障害物のない広い環境を配備し、外界との接触等による外乱はないものと仮定して調速行動実験を行う。図 6.15 については、学習開始のスタートボタン押下まではロボットは人間の手で支えられており、開始と同時に人間は手を離しそれ以降は人間によるロボットへの接触・干渉や操縦指令などの介入はないとする。

式 6.30 の PID ゲインの学習率は v, ω それぞれで表 6.6 に示す値に設定する。

表 6.6: 式 6.30 における PID ゲインの学習率 $\text{diag}(\eta_P, \eta_I, \eta_D)$ の設定

v に対する学習率	ω に対する学習率
$\text{diag}(10^{-7}, 10^{-7}, 10^{-7})$	$\text{diag}(10^{-9}, 10^{-9}, 10^{-9})$

図 6.15 -(a1)-(a6)，図 6.16 -(a1)-(a4) は、各時刻での調速行動の学習中におけるキャプチャを表しており、図 6.16 -(a0) は学習前の乗り込み動作のキャプチャを表す。図 6.16 の学習中の時刻については、乗り込み動作の完了後を $t = 0.0[\text{s}]$ とする。図 6.15 -(b1),(b2) および図 6.16 -(b1),(b2) が制御実行下における PID ゲイン全 6 パラメータのオンラインチューニングの推移であり、(b1) が速度 v の制御系、(b2) が角速度 ω の制御系に対応する。ここで、PID ゲイン変化のグラフの縦軸目盛は $\times 10^{-3}$ としていることに留意する。

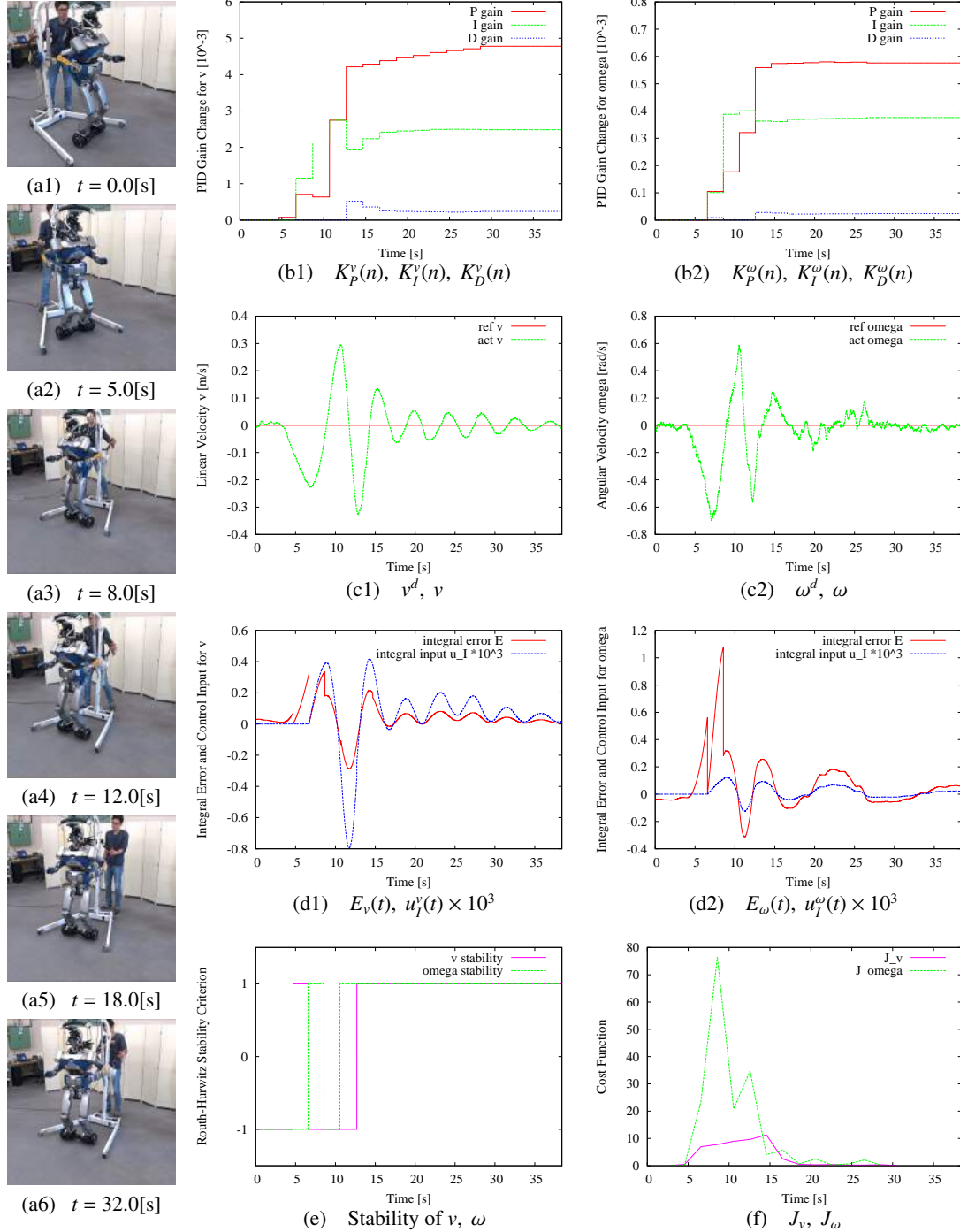


図 6.15: Experimental Results for Speed Governing Behavior by Life-Sized Humanoid Robot HRP2-JSK without Riding Motion

(a1)-(a6) show the experimental snapshots. (b1),(b2) show the change of updated PID gains [$\times 10^{-3}$] for v, ω (the red solid lines represent $K_p^v(n), K_p^\omega(n)$ and the green dashed lines represent $K_I^v(n), K_I^\omega(n)$, and the blue dashed lines represent $K_D^v(n), K_D^\omega(n)$). (c1),(c2) show the plots of linear velocity v [m/s] and angular velocity ω [rad/s] (the red solid lines represent v_d, ω_d , and the green dashed lines represent actual measured v, ω). (d1),(d2) show the plots of integrated value of control error and its control input (the red solid lines represent $E_v(t), E_\omega(t)$, and the blue dashed lines represent $u_I^v(t), u_I^\omega(t) \times 10^3$). (e),(f) show the change of Routh-Hurwitz stability and cost function (magenta solid lines: v , green dashed lines: ω).

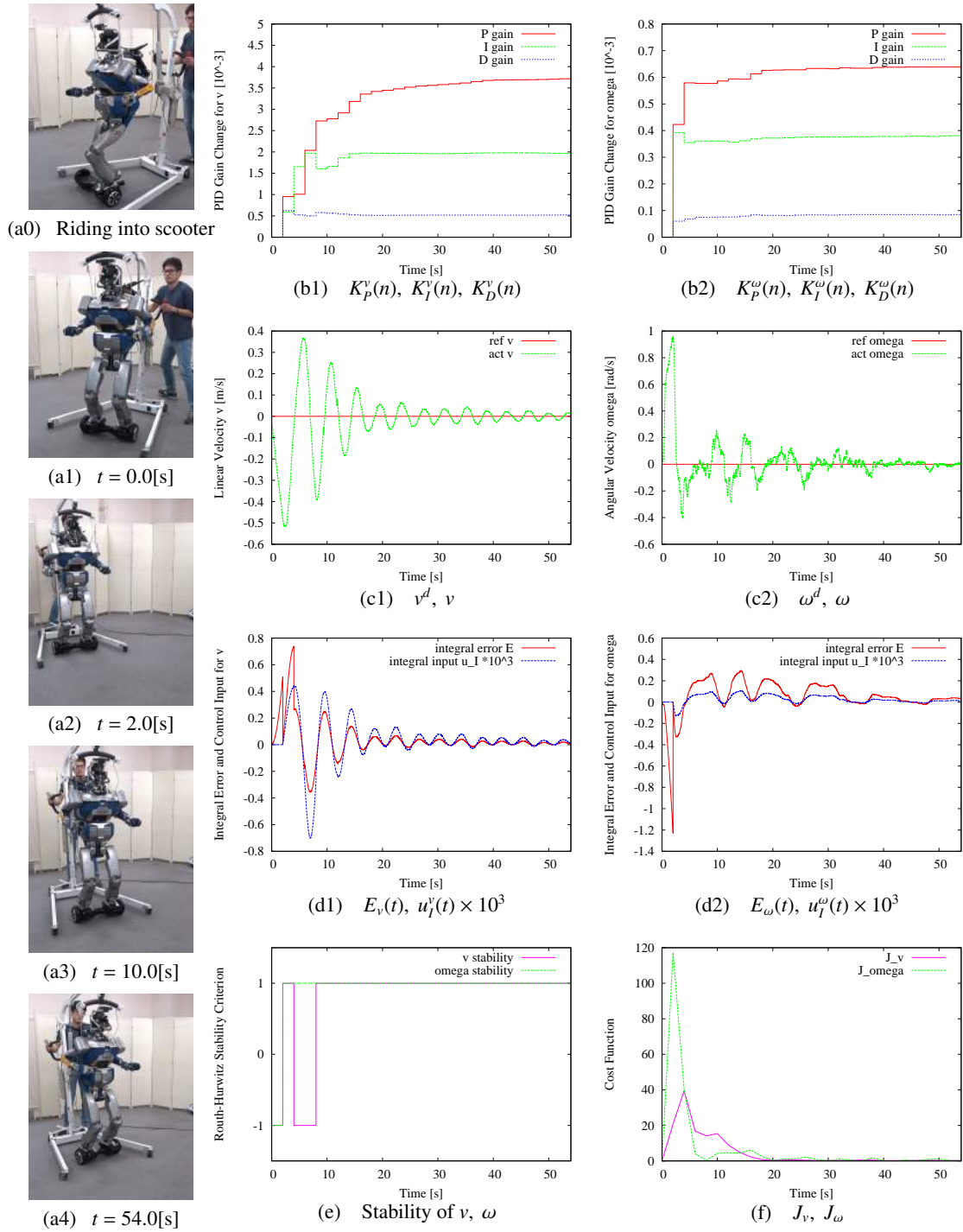


図 6.16: Experimental Results for Speed Governing Behavior by Life-Sized Humanoid Robot HRP2-JSK with Riding Motion

(a1)-(a6) show the experimental snapshots. (b1),(b2) show the change of updated PID gains $[\times 10^{-3}]$ for v, ω (the red solid lines represent $K_p^v(n), K_p^\omega(n)$ and the green dashed lines represent $K_I^v(n), K_I^\omega(n)$, and the blue dashed lines represent $K_D^v(n), K_D^\omega(n)$). (c1),(c2) show the plots of linear velocity v [m/s] and angular velocity ω [rad/s] (the red solid lines represent v_d, ω_d , and the green dashed lines represent actual measured v, ω). (d1),(d2) show the plots of integrated value of control error and its control input (the red solid lines represent $E_v(t), E_\omega(t)$, and the blue dashed lines represent $u_I^v(t), u_I^\omega(t) \times 10^3$). (e),(f) show the change of Routh-Hurwitz stability and cost function (magenta solid lines: v , green dashed lines: ω).

各 PID ゲインは初期値を 0 とした状態から学習を開始させ、図 6.15 は約 38[s]、図 6.16 は約 54[s] で式 6.32 の学習制御の終了条件を満たして収束した。図 6.16 は図 6.15 と比べて、人間によるサポート無しで乗り込んだ状態からの学習となるため、収束には多少の時間がかかっていることが推察される。また、学習結果から、D ゲインに比べて P ゲイン、I ゲインが制御系として支配的であることが分かる。

このときの調速の対象である速度 v と角速度 ω の目標値 $v^d = 0, \omega^d = 0$ への追従を図 6.15 -(c1),(c2)、図 6.16 -(c1),(c2) にそれぞれ示す。実験キャプチャ (図 6.15 -(a5),(a6)、図 6.16 -(a3),(a4))、およびグラフ (図 6.15 -(c1),(c2) の 35[s] 付近、図 6.16 -(c1),(c2) の 50[s] 付近) での目標値追従より調速行動が達成されていることが確認できる。

また、偏差の時間積分値 $E_*(t)$ と I 動作の制御入力 $u_I^*(t)$ の変化を、 v, ω 各々で図 6.15 -(d1),(d2)、図 6.16 -(d1),(d2) に示し、前回値補正積分器の効果を検証する。(d1),(d2) において赤線が偏差の時間積分値、青線が I 動作の制御入力 ($\times 10^3$) を表している。図 6.15 の 5[s]-8[s]、図 6.16 の 2[s] や 4[s] では、I ゲインの増加更新 $K_I^*(n+1) > K_I^*(n)$ により前回値補正積分器が SW1 へ切り替わり、式 6.37 に基づく前回積分値の補正が実行されていることが $E_v(t), E_\omega(t)$ のグラフ (赤線) の急峻な変化として観察できる。一方、それに対応する I 動作の制御入力 $u_I^v(t), u_I^\omega(t)$ は滑らかに変化しており、制御実行中の PID ゲインの変動 (b1),(b2) に影響されず発散なく制御入力を与えることに成功している。低い I ゲインのため補正前の積分値 $E_\omega(t)$ が修正しきれずに増大しており、そこに高い I ゲインの更新が掛けられて発散してしまう状況を、前回積分値を $\frac{K_I^*(n)}{K_I^*(n+1)}$ 倍して小さく補正することで未然に防いでいる。20[s] 以降など学習が収束に向かい I ゲインがほとんど更新されなくなると、前回積分値の補正処理もほぼ無くなり通常の偏差の積分演算 (SW0) となる。

最後に、調速行動の学習中における v, ω それぞれのフィードバック制御系の内部安定性を図 6.15 -(e)、図 6.16 -(e) に時間変化として表し、このときの IFT の評価関数 J_v, J_ω (式 6.17) の変化を図 6.15 -(f)、図 6.16 -(f) で観察する。内部安定性はプラント $P_*(s) = \frac{K_*}{s(1+T_*s)}$ とコントローラ $C_*(s) = K_p^*(n) + \frac{K_I^*(n)}{s} + K_D^*(n)s$ から定められる以下の特性多項式 $\phi(s)$ について、 $\phi(s) = 0$ のすべての根の実部が負であるか否かから判定する。

$$\phi(s) := T_*s^3 + (1 + K_*K_D^*(n))s^2 + K_*K_p^*(n)s + K_*K_I^*(n) \quad (6.40)$$

本稿では Routh-Hurwitz の安定判別法により PID ゲイン $K_p^*(n), K_I^*(n), K_D^*(n)$ が満たすべき内部安定性の条件式 6.41 を求めた上で、逐次更新される PID ゲインに対して内部安定の場合

は1, 内部安定でない場合は-1として(e)に結果を示した.

$$K_p^*(n) > 0, \quad K_I^*(n) > 0, \quad K_D^*(n) \geq 0, \quad K_p^*(n)(1 + K_*K_D^*(n)) - T_*K_I^*(n) > 0 \quad (6.41)$$

図 6.16 の結果では, v は 4[s]-8[s] で ω は 0[s]-2[s] において不安定な状態となり, この期間では評価関数 J_v, J_ω も各々高まって IFT の活性化に寄与していることが図 6.16 -(f) から確認できる. 学習が収束するにつれて安定な状態が継続して, 評価関数の値もほぼ 0 に落ち着いている. この状態変化は図 6.15 の結果からも同様に観察できる.

6.4.2 Iゲイン比に飽和制約を設けた方法と前回値補正積分器の比較

式 6.36 で示した I ゲイン更新に伴う制御入力の増幅率 (I ゲイン比) $\frac{K_I^*(n+1)}{K_I^*(n)}$ に飽和制約を設定することで制御入力の発散を抑える方法と, 提案した前回値補正積分器による手法について本項で効果を比較する. 前者は, 式 6.42 のように I 動作の制御入力 $u_I^*(t)$ の更新式に関して I ゲイン比 $\frac{K_I^*(n+1)}{K_I^*(n)}$ に飽和制約を与えることで, $K_I^*(n) \rightarrow K_I^*(n+1)$ の大きな変化に対して制御入力の発散を抑制する方法とする.

$$u_I^*(t + \Delta t) = \begin{cases} r_{\text{sat}}^* u_I^*(t) + \frac{\Delta t}{2} (e_*(t + \Delta t) + e_*(t)) K_I^*(n+1), & \frac{K_I^*(n+1)}{K_I^*(n)} > r_{\text{sat}}^* \\ \frac{K_I^*(n+1)}{K_I^*(n)} u_I^*(t) + \frac{\Delta t}{2} (e_*(t + \Delta t) + e_*(t)) K_I^*(n+1), & \frac{K_I^*(n+1)}{K_I^*(n)} \leq r_{\text{sat}}^* \end{cases} \quad (6.42)$$

r_{sat}^* ($*$ = v, ω) は v, ω に対する I ゲイン比の飽和制約の閾値である. 図 6.16 の結果を利用して, v の偏差の時間積分値および前回値補正積分器を適用した場合の I 動作の制御入力のグラフ (I ゲイン増幅が顕著な図 6.16 -(d1) での 0[s]-6[s]) に加えて, 式 6.42 による飽和制約 $r_{\text{sat}}^v = 1.5$ を設けた場合の制御入力をプロットした結果が図 6.17 である.

図 6.17 における緑線が飽和制約を与えた場合の I 動作の制御入力を表す. 飽和制約を与えた場合 (緑線) と前回値補正積分器を適用した場合 (青線) との大きな違いは 4[s] におけるスイッチングから観察でき, 飽和制約を設けた場合 (緑線) は瞬間的な変化を伴い制御入力が増加している. 一方で前回値補正積分器を適用した場合 (青線) は, I ゲイン切替時に偏差の時間積分 (赤線) の前回値を補正する (ここでは減少させる) ことで, I 動作の制御入力の突発的な立ち上がりを防いで滑らかな制御入力変化を実現している.

これらの観察事項は, $u_I^v(t)$ の離散時間での傾き $(u_I^v(t + \Delta t) - u_I^v(t)) / \Delta t$ を計算することで定量的に示すことができる. 飽和制約を与えた場合, $t = 4[s]$ に対する $(u_I^v(t + \Delta t) - u_I^v(t)) / \Delta t$ は式 6.42 の $\frac{K_I^v(n+1)}{K_I^v(n)} > r_{\text{sat}}^v$ より式 6.43 で表される. 一方で前回値補正積分器を適用した場合は,

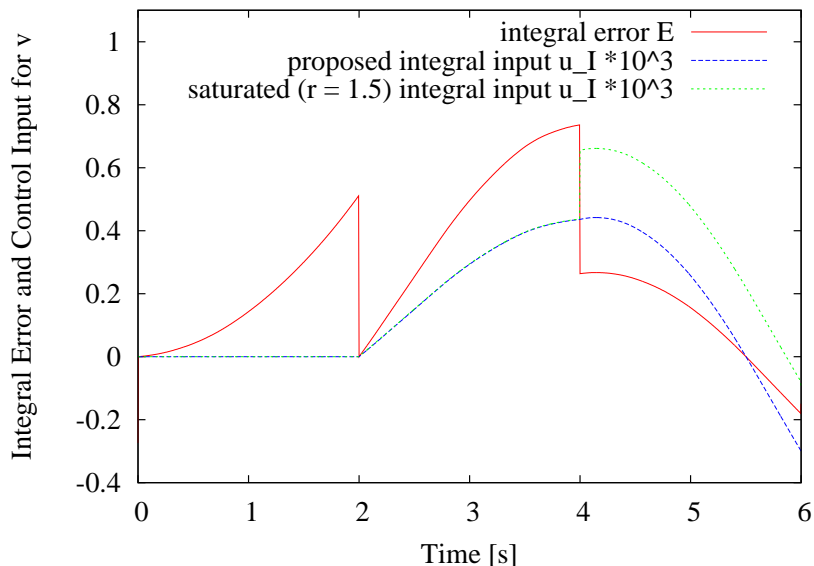


図 6.17: Comparison between the Proposed Integral Input for v (Blue Dashed Line) and the Saturated ($r_{\text{sat}}^v = 1.5$) Integral Input for v (Green Dashed Line)

The red solid line represents the plot of integrated value of control error $E_v(t)$ by applying proposed integrator.

式 6.38 より式 6.44 となる.

$$\frac{u_I^v(t + \Delta t) - u_I^v(t)}{\Delta t} = \frac{e_v(t + \Delta t) + e_v(t)}{2} K_I^v(n + 1) + \frac{r_{\text{sat}}^v - 1}{\Delta t} u_I^v(t) \quad (6.43)$$

$$\frac{u_I^v(t + \Delta t) - u_I^v(t)}{\Delta t} = \frac{e_v(t + \Delta t) + e_v(t)}{2} K_I^v(n + 1) \quad (6.44)$$

以上から、飽和制約を与えた場合 (式 6.43) は、前回値補正積分器を適用した場合 (式 6.44) に比べて $(r_{\text{sat}}^v - 1)u_I^v(t)/\Delta t$ だけ余分な傾きがかかっていることが確認できる。 $(r_{\text{sat}}^v - 1)u_I^v(t)/\Delta t$ の項が I 動作の制御入力の突発的な増加の原因であり、 $(r_{\text{sat}}^v - 1)u_I^v(t)/\Delta t = 0$ とするためには $r_{\text{sat}}^v = 1$ であればよく、このとき式 6.39 の前回値補正積分器の作動条件 $\frac{K_I^v(n+1)}{K_I^v(n)} > 1$ (SW1) と等価となる。

6.5 本章のまとめ

本章では、搭乗環境に属するヒューマノイドが搭乗型機器の移動性を把握するための運動への適応行動戦略を述べた。ヒューマノイドが搭乗型機器の扱いを獲得することの意味は、その搭乗型機器がもつ運動性能をヒューマノイド自身が知ることであり、実装された制御系

に対してどのようなプロパティ（どのようなゲイン）を与えるべきかを試行錯誤すること
に等しい。

本章においては、ここでの“実装された制御系”として、搭乗型機器上のヒューマノイド
が自身のスピードを把握することを目的とした調速行動制御系を新たに提案した。調速行動
制御系がもつ大きな特徴として、従来のバランス安定化制御などで扱われている接触対象
（本研究においては搭乗型機器相対）を基準とした議論から、接触対象の系外（本研究にお
いては搭乗型機器外の世界座標）を基準とした謂わば飛行ロボットや UAV をはじめとした
浮遊系としてヒューマノイドの状態・制御を扱えるように拡張した点にある。その拡張を補
助する機能構成として、ヒューマノイドがもつ視覚情報や慣性情報を把握するセンシング技
術が調速行動制御系には備わっている。

また、ヒューマノイドが搭乗した環境に適応的であるためには、新たな制御系の構築に加
えて、導入した制御系に対して“どのようなプロパティ（どのようなゲイン）を与えるべき
かを試行錯誤すること”へのアプローチについても解を与える必要がある。そのアプローチ
として制御系のゲイン学習は必須要素となる。

本章のゲイン学習方策では、第5章で提案した即時性を優先した SGD-based open-loop
Learning Control (SLC) とは対照的に、制御系のフィードバックも考慮に入れた closed-loop
でのミニバッチ処理形式のオンライン学習制御法として Mini-batch-based closed-loop Learning
Control (MLC) を提案した。MLC はヒューマノイドが試行錯誤する際に、新たなゲインを決
める働きをもつ脳に相当する部分であり、その決定を実際に稼働中の制御系に伝えて適用す
るのをサポートする働きが前回値補正積分器である。

本章では、上記の MLC と前回値補正積分器を統合することによりオンライン学習制御系
を構成し、調速行動制御系の内部の PID コントローラに対して適応的なゲイン学習を行う
ことで、平行二輪電動スクーターの搭乗環境の運動に適応したヒューマノイドによる調速行
動を実証した。

第7章

オンライン探索指標の 並行処理と切替処理による行動実現

7.1 はじめに

本章では、第2章で示した搭乗型機器の運転操作と適応行動に対するアプローチに基づき、第3章～第6章で提案した軌道修正とゲイン学習において共通戦略となるオンライン探索指標の概念を用いた複合的な処理方法（“並行処理”と“切替処理”）について述べる。第3章～第6章において整理したオンライン探索指標に対して、これら2つの処理を適用し実行させることでヒューマノイドによる搭乗型機器への統合的な行動実現を示す。

7.2節では、異なるオンライン探索指標に対しての2通りの処理方法（“並行処理”と“切替処理”）を述べ、第2章のアプローチをもとにした本章の立ち位置を明確化する。7.3節では、これまでの議論で整理した一般化更新則（式2.13）に対するオンライン探索指標をはじめとした各パラメータの対応をもとに、オンライン探索指標の並行処理による行動実現を実証する。7.4節では、7.3節と同様にオンライン探索指標まわりのパラメータを整理し、その対応をもとにした切替処理による行動実現を実証する。

7.2 オンライン探索指標の並行処理と切替処理

これまでに論じてきた第3章、第4章、第5章、第6章では、各々で提案した軌道修正とゲイン学習の適用戦略に関して、第2章で述べた全体の共通軸となるオンライン探索指標の考え方をを用いることで、ヒューマノイドが搭乗型機器を運転操作し、適応行動を実現することへの理解を深めた。本章で論じる内容は、これらのオンライン探索指標を利用した軌道修正やゲイン学習を統合的に扱った同時的または断続的な処理方法の考え方である。本研究ではこれらの処理を“オンライン探索指標の並行処理”と“オンライン探索指標の切替処理”と定義する。

第2章で確立した搭乗型機器への運転操作と適応行動における集合と写像の考えを利用したアプローチに基づき本章の立ち位置を理解すると、図7.1の橙色で示された共通部分の領域に属する。この領域内に存在する行動課題をオンライン探索指標の並行処理と切替処理の適用により解くことで、搭乗型機器に対する複合的な行動実現を目指す。

“オンライン探索指標の並行処理”と“オンライン探索指標の切替処理”はそれぞれ以下に示すような処理手法である。

オンライン探索指標の並行処理：

異なるオンライン探索指標同士を共通の制御システム/学習システム上で共存させた同

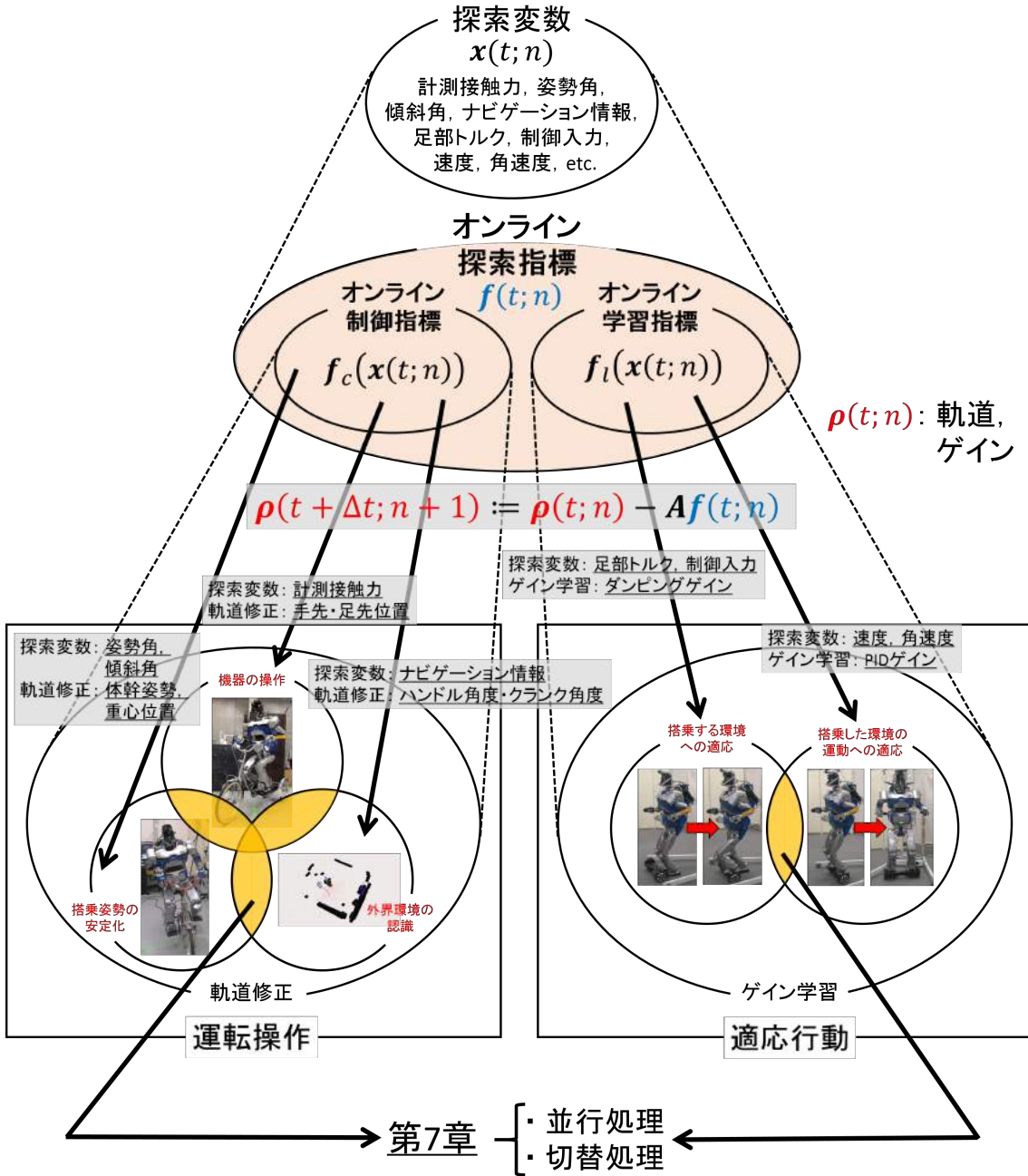


図 7.1: オンライン探索指標を基にして解かれる搭乗型機器の運転操作と適応行動における各課題と本章が属する立ち位置の関係
本章では、提案したオンライン探索指標の並行処理と切替処理の手法を利用して、橙色の共通部分に存在する行動課題に対して解を与える。

時的に実行される処理手法

オンライン探索指標の切替処理：

異なるオンライン探索指標同士が個々の制御システム/学習システム間で切替遷移される断続的な処理手法

オンライン探索指標の並行処理では、異なって定義された別々のオンライン探索指標が制御システム/学習システム上で共有し合い、異なる軌道修正/ゲイン学習の処理を同時実行する行動実現を目的とする。具体的に閉ループ系で構成されたシステム上においては、別々のオンライン探索指標同士がループ処理内で共存し作用し合う働きを有している。

一方で、オンライン探索指標の切替処理では、異なって定義された別々のオンライン探索指標は共有することなくそれぞれ独立した制御システム/学習システムをもっているが、構成されたシステム間でオンライン探索指標が行き来する切替遷移の働きを有しており、これにより断続的に機能を切り替えた互いに干渉のない行動実現を可能としている。

本研究ではこれら2つのオンライン探索指標に対する処理手法を、ヒューマノイドにおける搭乗型機器の運転操作および適応行動に実装し、複合的な行動実現へと発展させる。

7.3 並行処理による行動実現

7.3.1 搭乗型機器の操作と実時間並行処理型姿勢制御に基づく並行処理実験

第3章で述べた搭乗型機器の操作に対する軌道修正法と、4.3節の搭乗姿勢の安定化に対する軌道修正法として提案した実時間並行処理型姿勢制御(4.3.3項)を共存させた行動実現を示す。本行動実現においては、搭乗型機器の操作における一般化更新則(式2.13)とのパラメータ対応(表3.3)、および実時間並行処理型姿勢制御における一般化更新則とのパラメータ対応(表4.4)に基づいた、表7.1に示す2つの処理系のオンライン探索指標 $f(t;n)$ 、更新される軌道 $p(t;n)$ 、一般化更新則の重み A に対する並行処理を実行する。

並行処理による行動実現を実証した実験として、図7.2に起伏を設けた床面での三輪車操作と運転操作中のヒューマノイドの自己姿勢を並行して安定化させる行動処理実験を示す。図7.2の運転実験環境では、走行路面に図7.2-(1)に示されるような約10[mm]厚の段差が2つ存在しており、搭乗型機器である三輪車の前輪・後輪が段差に乗り上げることによる車体傾斜を、ヒューマノイドが運転行動と同時に姿勢安定化を図る目的で設定している。割り込

表 7.1: 搭乗型機器の操作と実時間並行処理型姿勢制御の並行処理において対応する各パラメータ

並行処理	搭乗型機器の操作	実時間並行処理型姿勢制御
オンライン探索指標 $f(t; n)$	$\text{sgn}[\rho(t; n), f^{\text{msr}}(t)] (f^{\text{msr}}(t) - f^{\text{trg}}(t))$	$\text{sgn}[\rho(t; n), \phi_C(t)] (\phi_C(t) - \phi_C^d(t))$
更新される軌道 $\rho(t; n)$	$\xi^{\text{ref}}(t)$	$\phi_C^d(t)$
一般化更新則の重み A	$(M_v s^2 + D_v s + K_v)^{-1}$	$k_C T_C / (1 + T_C s)$

み処理型座り直し制御のように、運転操作の途中で安定化行動を実行するために操作を中断する処理は含まれておらず、表 7.1 に示したような 2 つのオンライン探索指標を同時に処理する共存した軌道修正により、搭乗型機器の操作と搭乗姿勢の安定化を並列実行している。

図 7.2 -(2) のように運転操作中の三輪車は最初の段差に後輪を乗り上げて車体が Roll まわりに傾斜するが、傾斜した搭乗環境に対してヒューマノイドがバランスを保つよう適応的に搭乗姿勢の安定化制御を実行している (図 7.2 -(3))。その後、段差を乗り越えて平坦な床面へと路面環境が変化 (図 7.2 -(4)) すると、その変化に応じてヒューマノイドは再び搭乗姿勢を修正しながら、次の段差に向かって運転操作を続行する。図 7.2 -(5) のように前輪が乗り上げた後、図 7.2 -(7) のように後輪が乗り上げる振動的な搭乗環境変化に対しても、実時間並行処理型姿勢制御の貢献により運転操作を途中で中断することなく転倒せずに移動可能なことを実証した。

以上のように搭乗環境に属するヒューマノイドは、表 7.1 のオンライン探索指標で示される力覚 $f^{\text{msr}}(t)$ や身体の慣性 $\phi_C(t)$ を複合的にセンシングしながら、自身がつまみピュレータのエンドエフェクタの軌道 $\xi^{\text{ref}}(t)$ および自身の体幹の軌道 $\phi_C^d(t)$ を並行して逐次修正することで、運転操作とそれに伴うバランス維持への対処を同時に実行できることを本行動実験の結果より得た。



(1)

(2) Rear Wheel onto Step-1



(3) Stabilize Roll Inclination

(4) Return Flat Floor



(5) Front Wheel onto Step-2

(6)



(7) Rear Wheel onto Step-2

(8) Stabilize Roll Inclination

図 7.2: Simultaneous Behavior of Vehicle Manipulation and Stabilizing Balance on the Rough Floor

7.3.2 外界環境に対する認識と搭乗型機器への操作量逐次修正に基づく並行処理実験

前項と同様に第3章で述べた搭乗型機器の操作に対する軌道修正法と、本項では4.4節の外界環境の認識に対する軌道修正法を取り入れることで共存させた並行処理に基づく行動実現を示す。本行動実験における並行処理は、表7.2に示すオンライン探索指標 $f(t; n)$ 、更新される軌道 $\rho(t; n)$ 、一般化更新則の重み A を組み合わせた処理系で実行され、搭乗型機器の操作に対するパラメータ対応は表7.1と同様である。

表7.2: 搭乗型機器の操作と外界環境に対する認識行動の並行処理において対応する各パラメータ

並行処理	搭乗型機器の操作	外界環境に対する認識行動
オンライン探索指標 $f(t; n)$	$\text{sgn}[\rho(t; n), f^{\text{msr}}(t)](f^{\text{msr}}(t) - f^{\text{trg}}(t))$	$\text{sgn}\left[\rho(t; n), \begin{pmatrix} \dot{\theta}^{\text{ref}}(t) \\ \dot{x}^{\text{ref}}(t) \end{pmatrix}\right] \left(\begin{pmatrix} \dot{\theta}^{\text{ref}}(t) \\ \dot{x}^{\text{ref}}(t) \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right)$
更新される軌道 $\rho(t; n)$	$\xi^{\text{ref}}(t)$	$\begin{pmatrix} \theta_h^{\text{ref}}(t) \\ \theta_c^{\text{ref}}(t) \end{pmatrix}$
一般化更新則の重み A	$(M_v s^2 + D_v s + K_v)^{-1}$	$\begin{pmatrix} L/\delta \dot{x}_*^{\text{ref}} & 0 \\ 0 & 2\pi\Delta t/D_c \end{pmatrix}$

外界環境に対する認識行動に関しての上記パラメータは、表4.6で示した大型障害環境に対する視覚ナビゲーション情報に基づいたパラメータ対応を表しており、本行動実験では外界環境として大型障害環境となる障害物（障壁）を設定した、ロボットの視覚情報を用いた環境把握・認識と搭乗型機器への操作量の逐次修正による並行処理の実現を目的とする。大型障害環境に対する認識行動戦略は4.4.2項で論じた手法を適用するものとして、本項ではヒューマノイドの自己位置情報も同時に逐次更新することで図4.14や図4.17で構築したclosed-loopな処理系における循環性を高める。

本行動実験を対象とした全体のシステムを図7.3に示す。図7.3は、第4章で提案した認識行動処理系（図4.14）および認識行動システム（図4.17）を、搭乗型機器の操作と外界環境の認識を統合した並行処理実験に当てはめたシステムを表している。並行処理実験では、人間が指示した目的地点までの障害環境（本実験では障壁）を考慮した三輪車の逐次的な操作修正に基づく自律的な回避適応行動を実証する。図7.3の特徴として、先に述べた提案システム（図4.14や図4.17）と比較して処理系のclosed-loopの構成が新たに増えており、

ヒューマノイドの自己位置推定の処理ブロックを介して自己位置の情報も closed-loop とし て並行処理系に含まれる。

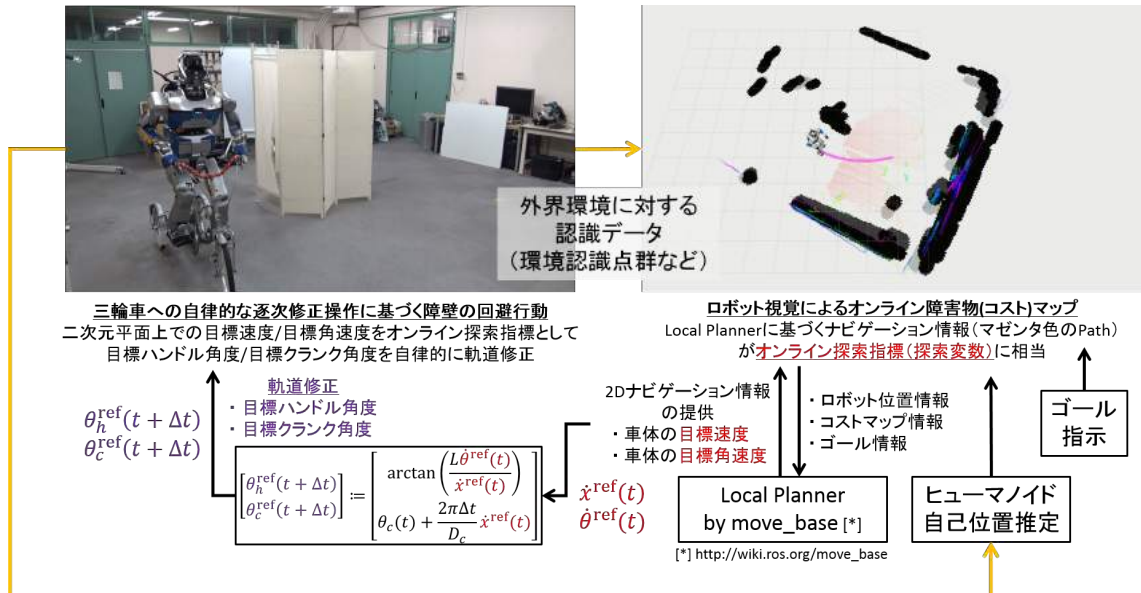


図 7.3: 外界環境に対する認識と搭乗型機器への操作量逐次修正に基づく並行処理実現システム

ヒューマノイドの自己位置情報に基づく搭乗型機器の移動成分分析

図 7.3 のシステムで新たに加えられたヒューマノイドの自己位置推定による逐次的なループに関して、ヒューマノイドの位置推定情報を利用した搭乗型機器の移動成分について分析する。一般的に移動ロボットの研究では Odometry の考えが用いられる。ヒューマノイドにおける搭乗型機器の運転操作では、主に Wheel Odometry と Visual Odometry に大別して分析し、またこれら 2 つの Odometry を相補する SLAM (Simultaneous Localization and Mapping) システムから推定される SLAM Odometry についても図 7.4 で分析を行う。図 7.4 が表している各 Odometry は以下の通りである。

Ground Truth Odometry:

移動成分の真値 (ヒューマノイド頭部に搭載されたステレオカメラからみた外界の checkerboard の位置・姿勢から計測)

Wheel Odometry:

搭乗型機器（本実験では三輪車）の移動モデル（ハンドル角度や車輪の回転数から推定）

Visual Odometry:

ヒューマノイド頭部に搭載されたステレオカメラおよびLRF（Laser Range Finder）によるビジョン情報・距離情報に基づく移動量

SLAM Odometry:

Wheel Odometry と Visual Odometry を相補する SLAM システム [64][190] から推定された Odometry

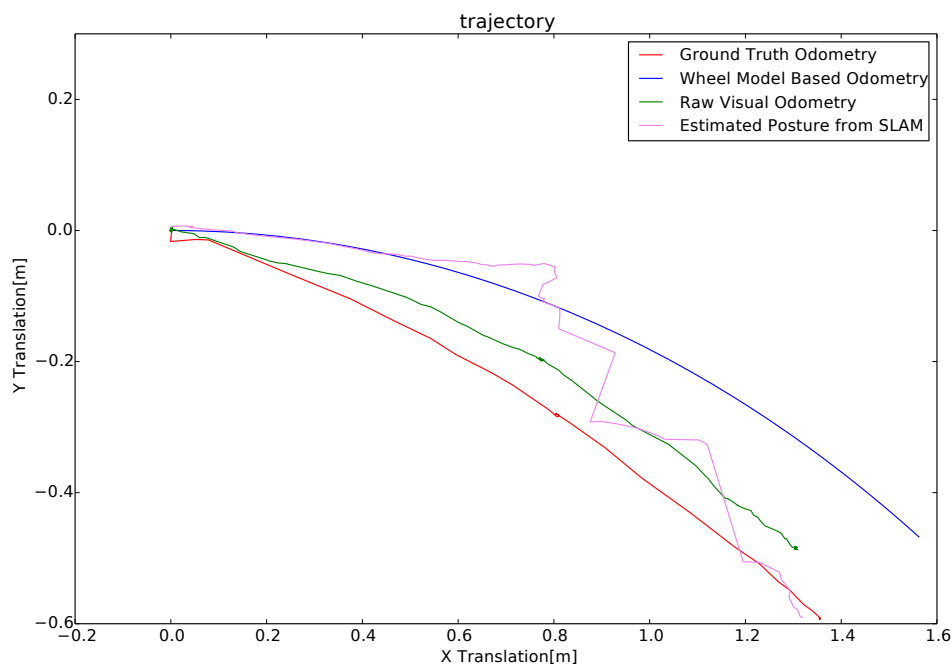


図 7.4: Odometry Analysis for Humanoid with Tricycle

In this movement on X-Y plane, humanoid robot drives and stops tricycle at $-25[\text{deg}]$ handle angle.

SLAM Odometry の推定法は文献 [64][190] で提案されている手法を利用し、最終的に SLAM Odometry を基に図 7.3 の自己位置情報の推定処理のループを構成する。図 7.4 は、等身大ヒューマノイドロボット HRP2-JSK [110] による三輪車操作時の Ground Truth Odometry, Wheel Odometry, Visual Odometry, SLAM Odometry の各分析結果であり、ハンドル角度

を $-25[\text{deg}]$ として二次元平面（X-Y平面）上で走行と停止の運動を行っている。図7.4の結果から、SLAM Odometryの特徴として次のことが観察される。

- 走行初期（ $(X, Y) = (0[\text{m}], 0[\text{m}])$ 付近）においては Wheel Odometry の値を信頼している。
- Visual Odometry から乖離するにつれて、Visual Odometry の結果を信頼して相補する（ $X = 0.8[\text{m}]$ 付近）。
- 最終地点（ $(X, Y) = (1.3[\text{m}], -0.6[\text{m}])$ 付近）では Ground Truth Odometry に最も近い位置に補正されている。

SLAM Odometry に基づく搭乗型機器への操作量逐次修正による障害物回避行動

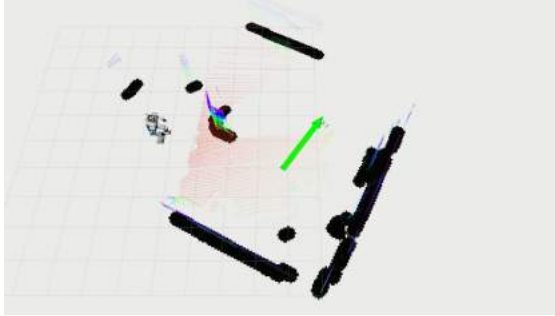
SLAM Odometry を図7.3の自己位置推定の処理系に導入し、ヒューマノイドおよび搭乗型機器における自己位置情報をリアルタイムに更新・反映させた自律的な搭乗型機器への逐次修正操作に基づく障壁の回避行動実験を図7.5に示す。図7.5-(a1)-(a4)は等身大ヒューマノイドロボット HRP2-JSK [110] による外界環境に存在する障害物（障壁）を回避した自律的な運転操作行動のキャプチャであり、図7.5-(b1)-(b4)がロボット視覚によるオンラインでの障害物情報を反映した障害物マップ（コストマップ）を表す。

図7.5-(b1)のように最初に人間がモニター上の Viewer システム（図4.18の操作インターフェース）から緑のマーカーで示されたゴールを指定する。リアルタイムに更新されるロボット位置情報、視覚認識によるコストマップ情報、人間が与えたゴール情報を基にして、図7.5-(b2)のマゼンタ色の Path で示されるように、move_base [79] の Local Planner から計画された二次元平面上での車体のナビゲーション情報（目標速度 $\dot{x}^{\text{ref}}(t)$ /目標角速度 $\dot{\theta}^{\text{ref}}(t)$ ）が提供される。

ナビゲーション情報 $(\dot{\theta}^{\text{ref}}(t) \ \dot{x}^{\text{ref}}(t))^{\text{T}}$ を探索変数 $\mathbf{x}(t; n)$ とした表7.2に従ったオンライン探索指標の生成により、搭乗型機器へのハンドル角度 $\theta_h^{\text{ref}}(t)$ /クランク角度 $\theta_c^{\text{ref}}(t)$ の目標操作量に対する逐次的な軌道修正が実行される。軌道修正から決定された目標ハンドル角度と目標クランク角度をロボット実機に指令し、表7.2の搭乗型機器の操作に対する軌道修正を実行することで実環境に作用した障害物回避行動を実現している。実環境の変化に伴い、外界環境に対する認識データ（環境認識点群など）および SLAM Odometry によるロボット自己位置、が更新されることで closed-loop に並行処理が実現される。



(a1)



(b1) Input Goal (Green Marker) on Viewer



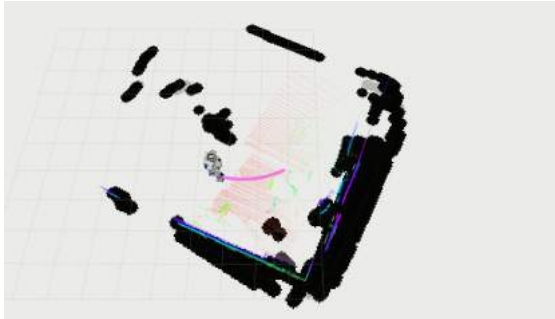
(a2) Modification for Steering and Pedaling



(b2) Navigation



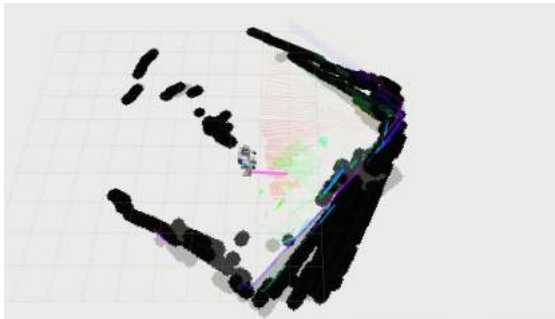
(a3)



(b3)



(a4) Toward Goal



(b4) Navigation to Converge

図 7.5: Simultaneous Behavior of Vehicle Manipulation and Recognition for Outside Obstacle Environment

(a1)-(a4) show the experimental snapshots of avoidance behavior. (b1)-(b4) show the viewer screenshots on monitor.

最終的に本行動実験では操作量逐次修正により図 7.5 -(a4),(b4) に示す通り，ゴール地点に向かってナビゲーション情報が収束していく様子が確認できる．この観察内容は表 7.2 で示したオンライン探索指標の収束として考えることができ，収束に従って更新される軌道 $(\theta_h^{\text{ref}}(t) \theta_c^{\text{ref}}(t))^T$ も零ベクトルへと逐次漸近し，自動的に搭乗型機器への操作を停止することが可能となっている．また，搭乗型機器の移動成分を把握した自律系を構築することで，搭乗型機器を扱うヒューマノイドが次に実行すべき操作行動とそれに対する操作量を逐次決定することを可能とした．

7.3.3 目標 ZMP の軌道修正と調速行動のゲイン学習に基づく並行処理実験

これまでに，第3章，第4章で述べてきた軌道修正の考え方と，第5章，第6章で述べてきたゲイン学習の考え方は，各々共存することなく独立した集合の概念として捉えてきた．本項では，これら2つの方策が共存するような条件の下での並行処理に基づいた行動実現を示す．これまでの議論では，所望とする各行動実現に対して人間の判断により，軌道修正法を適用するのが適切か，ゲイン学習法を適用するのが適切か，について熟考した上で方策の処理を行ってきている．本項で示す行動実現においてはどちらの方策が支配的であるかについても未知とした上で，軌道修正とゲイン学習が共存したシステムを構成することにより，人間が与えてきた既知の情報を可能な限り狭めていくことを目的とする．

本項の搭乗型機器に対する行動実現として，A.1.3 項に示すような一輪の電動車輪で駆動する電動一輪車 “Ninebot One S2” [191] に対するヒューマノイドの適応行動を検証する．適応行動実験においては，一輪車の Roll まわりの不安定性をハードウェアとして補う用途で図 A.8 -(II) に示すように左右に受動的な補助輪を設けている．適応行動で獲得した軌道/ゲインを利用することで，補助輪を外した電動一輪車の運転操作を実証する．

補助輪を有した電動一輪車上でのヒューマノイドの適応課題

ヒューマノイド実機の関節角度に搭乗姿勢の指令値を与え，AutoBalancer [85] を適用したオフラインでの安定化姿勢を求めた上で Linear Inverted Pendulum Tracking Controller [55] による安定化制御を導入する．この実機状態で補助輪を有した電動一輪車 “Ninebot One S2” [191] に等身大ヒューマノイドロボット HRP2-JSK [110] を搭乗させたキャプチャを図 7.6 に示す．搭乗型機器である “Ninebot One S2” 自体のバランスは，Roll 方向に対しては補助輪による

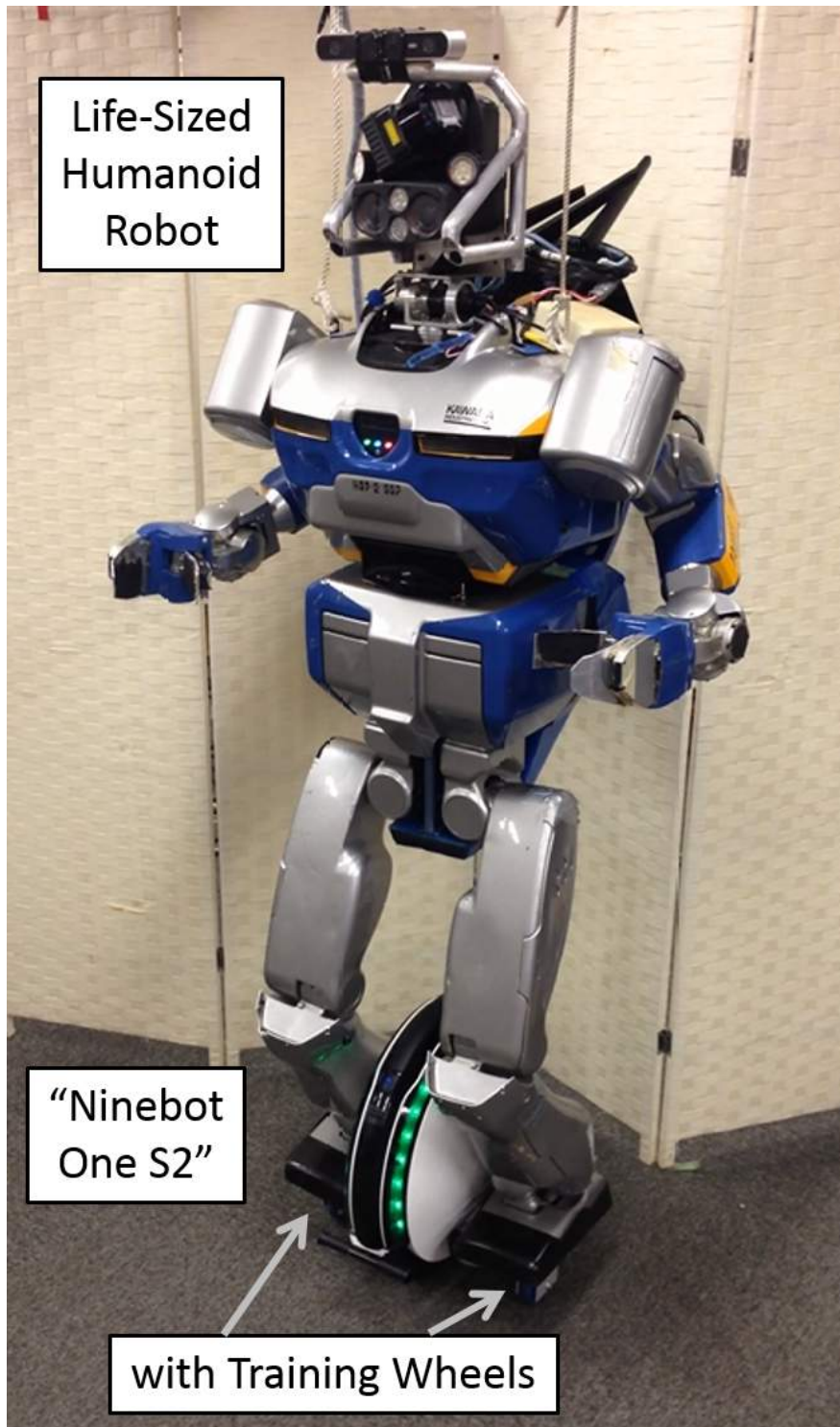


図 7.6: Overview of Life-Sized Humanoid Robot on “Ninebot One S2” with Training Wheels to Support Roll Balance

ハードウェアでの支持，Pitch 方向に対しては車体の傾斜に応じた車輪の電動制御によるバランスの支持機能を備えている。

“Ninebot One S2” 上においてヒューマノイドは AutoBalancer [85] や Linear Inverted Pendulum Tracking Controller [55] により安定化が実現されるが，図 7.7 のキャプチャに示すようにに搭乗するヒューマノイドのバランス安定化制御と “Ninebot One S2” がもつ Pitch 方向に対するバランス支持機能が共存する環境では速度の制御が達成されておらず，図 7.7 -(1) のように人手のサポートを外すと図 7.7 -(2),(3),(4) からも観察される通り，速度をゼロとするようなその場での定位が困難な課題をもつ。



図 7.7: Behavior of Life-Sized Humanoid Robot on “Ninebot One S2”

In this experiment, we apply the balance control (AutoBalancer [85] and Linear Inverted Pendulum Tracking Controller [55]) to humanoid robot with “Ninebot One S2” balance support.

軌道修正とゲイン学習の並行処理による调速行動実現

前述の “Ninebot One S2” 上でのヒューマノイドの適応課題を解くために，第 6 章で論じた搭乗環境の運動に適応する调速行動の実現を図る。第 6 章では平行二輪電動スクーターを搭

乗型機器例に挙げ、調速行動制御系に対するゲイン学習により調速行動を実証した。平行二輪電動スクーターの駆動特性は A.1.2 項で示されるように既知とした上で、搭乗するヒューマノイドにおける調速行動制御系を構成し PID ゲインに対するオンライン学習制御系を導入しているが、“Ninebot One S2”の機器例では平行二輪電動スクーターで見られるような足部ボードの傾斜量と駆動輪の関係が未知であり、また搭乗者の重心・ZMP の移動（重心・ZMP 軌道の修正）に依存した駆動特性となるかの判断が困難である。そこで、第6章で提案した調速行動制御系と共存する形でヒューマノイドの目標 ZMP に対する軌道修正を並行処理として実現させることで、軌道修正ならびにゲイン学習の貢献度合いを未知とした条件での調速行動を達成させる。

目標 ZMP の軌道修正を実行する系（軌道修正系と呼ぶ）と、図 6.6 で示したゲイン学習を実行する調速行動制御系の併合システムを図 7.8 のように設計する。また、図 7.8 で示す各パラメータを表 7.3 に従い定義する。軌道修正系の PID コントローラ $C_{traj}(s)$ ならびに調速行動制御系の PID コントローラ $C_v(s)$ は、共に Mini-batch-based closed-loop Learning Control (MLC) を利用した PID ゲインの適応的な学習により $K_p^{traj}(n), K_I^{traj}(n), K_D^{traj}(n)$ および $K_p^v(n), K_I^v(n), K_D^v(n)$ がオンラインでセットされる。図 7.8 の併合システムにおいては、Mini-batch-based closed-loop Learning Controller のブロックが軌道修正系と調速行動制御系のそれぞれの処理系の coupling としての役割を担っており並行処理実現を可能としている。これにより軌道修正系の制御入力 u_{traj} を与えることで、従来の Linear Inverted Pendulum Tracking Controller [55] 処理後の目標 ZMP p^{d*} から最終的に指令する目標 ZMP（最終目標 ZMP） p^{d**} への適応的な軌道修正が実現される。

$C_{traj}(s), C_v(s)$ の伝達関数は共に PID コントローラであることから共通に以下のように表される。

$$C_*(s) = K_p^*(n) + \frac{K_I^*(n)}{s} + K_D^*(n)s, \quad * = traj, v \quad (7.1)$$

目標 ZMP の軌道修正による更新式は

$$p^{d**} = p^{d*} + u_{traj} \quad (7.2)$$

で表され、計算機上では最終目標 ZMP の時間微分 $\dot{p}^{d**}(t)$ は次式で算出される。

$$\begin{aligned} \dot{p}^{d**}(t) &= \text{LPF}_{T_{traj}} \left\{ \frac{p^{d**}(t) - p^{d**}(t - \Delta t)}{\Delta t} \right\} \\ &= \text{LPF}_{T_{traj}} \left\{ \frac{u_{traj}(t)}{\Delta t} \right\} \end{aligned} \quad (7.3)$$

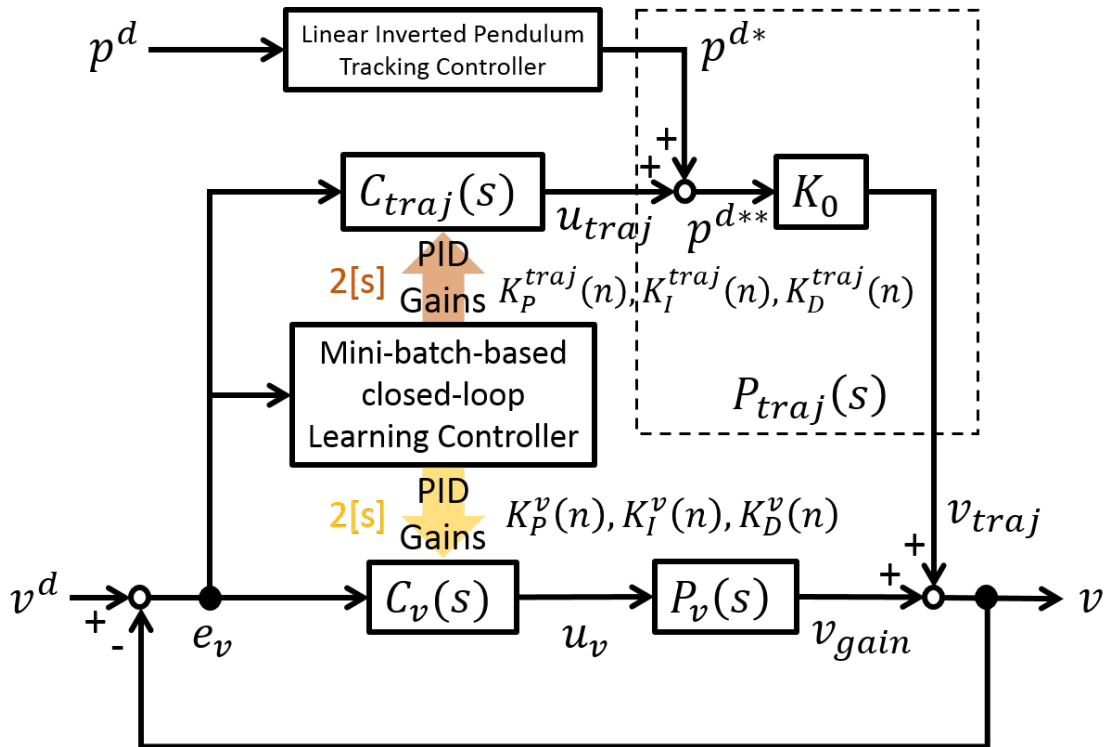


図 7.8: 軌道修正とゲイン学習の並行処理に基づく併合システム

ここで、 Δt は制御周期であり、時間微分によるノイズを除去するためのローパスフィルタ $LPF_{T_{traj}}\{\}$ (T_{traj} : 時定数) を施している. 式 7.3 の両辺をラプラス変換し、 p^{d**} について解くと以下となる.

$$p^{d**} = \frac{1}{\Delta t} \frac{1}{s(1 + T_{traj}s)} u_{traj} \quad (7.4)$$

“Ninebot One S2” における最終目標 ZMP p^{d**} と出力される並進速度 v_{traj} の関係は式 A.9 で与えられ、式 A.9 に式 7.4 を代入することで u_{traj} と v_{traj} の関係すなわち軌道修正系のプラントの伝達関数 $P_{traj}(s)$ は以下で記述される.

$$\begin{aligned} v_{traj} = P_{traj}(s)u_{traj} &= \frac{K_0}{\Delta t} \frac{1}{s(1 + T_{traj}s)} u_{traj} \\ &= \frac{K_{traj}}{s(1 + T_{traj}s)} u_{traj} \end{aligned} \quad (7.5)$$

ここで、定ゲインを $K_{traj} = K_0/\Delta t$ としてまとめた. 式 7.5 で得た $P_{traj}(s)$ は调速行動制御系のプラント式 6.13 と同一の形式であり、 $P_{traj}(s)$ および $P_v(s)$ に関しても共通に扱うことが

表 7.3: 軌道修正系と調速行動制御系の併合システムにおける各パラメータの定義

p^d	Linear Inverted Pendulum Tracking Controller [55] 処理前の目標 ZMP
p^{d*}	Linear Inverted Pendulum Tracking Controller [55] 処理後の目標 ZMP
p^{d**}	軌道修正後の最終目標 ZMP
v^d	並進速度の目標値
v	並進速度
e_v	偏差 $v^d - v$
$C_{traj}(s)$	軌道修正系のコントローラ
$C_v(s)$	調速行動制御系のコントローラ
$K_P^{traj}(n), K_I^{traj}(n), K_D^{traj}(n)$	$C_{traj}(s)$ の PID ゲイン
$K_P^v(n), K_I^v(n), K_D^v(n)$	$C_v(s)$ の PID ゲイン
u_{traj}	軌道修正系の制御入力 (修正量)
u_v	調速行動制御系の制御入力
$P_{traj}(s)$	軌道修正系のプラント
$P_v(s)$	調速行動制御系のプラント
v_{traj}	軌道修正により出力される並進速度
v_{gain}	調速行動制御系のゲイン学習により出力される並進速度
K_0	入力 p^{d**} に対する出力 v_{traj} の “Ninebot One S2” 固有のゲイン

可能である。したがって、6.3 節で述べたオンライン学習制御のアルゴリズムを共有して適用することができ、これに従い “Ninebot One S2” 上での軌道修正とゲイン学習の並行処理に基づく調速行動実験を行った。目標 ZMP の軌道修正と調速行動のゲイン学習の並行処理におけるオンライン探索指標 $f(t; n)$ ，更新される軌道 $\rho(t; n)$ ，更新されるゲイン $\rho(t; n)$ ，一般化更新則の重み A の組み合わせを表 7.4 に整理する。

$v^d = 0$ とする調速行動の実験結果を図 7.9 に示す。図 7.9 -(a1)-(a4) が実験時のキャプチャであり、図 7.9 -(b),(c) に示されるように調速行動中において軌道修正系のゲイン $K_P^{traj}(n), K_I^{traj}(n), K_D^{traj}(n)$ と調速行動制御系のゲイン $K_P^v(n), K_I^v(n), K_D^v(n)$ が並行処理として学習されている。ゲインの学習周期は第 6 章で示した調速行動実験と同様にそれぞれ 2[s] 毎の更新として設定している。図 7.9 -(d) が並進速度の目標値 $v^d = 0$ (“ref v”) および実際の計測された並進速度 v (“act v”) のプロットを表し、図 7.9 -(e) が偏差 e_v に基づく学習制御の評価関数 (式 6.17) の時間変化を表す。

図 7.9 -(a4) や図 7.9 -(d) の終端時刻から観察できるように並進速度 v は発散なく目標値 $v^d = 0$ へと漸近していることが分かる。また、学習制御において設定した評価関数も図 7.9 -(e) に示す通り、0 へと収束に成功している。0[s]-2[s] においては未学習の状態 (全ゲイン

表 7.4: 目標 ZMP の軌道修正と調速行動のゲイン学習の並行処理において対応する各パラメータ

並行処理	目標 ZMP の軌道修正	調速行動のゲイン学習
オンライン探索指標 $f(t; n)$	$-u_{traj}$	$\left(\frac{\partial J(v)}{\partial \rho(t; n)}\right)^T$
更新される軌道 $\rho(t; n)$	p^{d*}	—
更新されるゲイン $\rho(t; n)$	$K_P^{traj}(n), K_I^{traj}(n), K_D^{traj}(n)$	$K_P^v(n), K_I^v(n), K_D^v(n)$
一般化更新則の重み A	1	$\text{diag}(\eta_P, \eta_I, \eta_D)$

$K_P^{traj}(n), K_I^{traj}(n), K_D^{traj}(n), K_P^v(n), K_I^v(n), K_D^v(n)$ が 0) であり, この期間は並進速度 v および評価関数の発散が観察できるものの, その後の各 PID ゲインの学習・チューニングにより調速行動が達成されている. 軌道修正系のゲインにおいては $K_P^{traj}(n)$ が, 調速行動制御系のゲインにおいては $K_P^v(n), K_I^v(n)$ が支配的となる結果が得られた. また, “Ninebot One S2” を用いた本実験で獲得した調速行動制御系のゲイン変化 (図 7.9 -(c)) を, 平行二輪電動スクーターの場合の実験結果 (図 6.15 -(b1), 図 6.16 -(b1)) と比較すると, “Ninebot One S2” での $K_P^v(n), K_I^v(n), K_D^v(n)$ の学習値は平行二輪電動スクーターでの学習値の 10 分の 1 未満であり, その一方で軌道修正系のゲイン変化 (図 7.9 -(b)) は $K_P^{traj}(n)$ を中心に大きな値を獲得していることが分かる. 上記の観察事項から, “Ninebot One S2” 上におけるヒューマノイドの適応性は調速行動制御系のゲイン学習よりも, 目標 ZMP に対する軌道修正のほうがより支配的であり貢献度として大きいことが考察できる.

調速行動時における最終目標 ZMP p^{d**} の x 成分 (“fin ref ZMP x”) の軌道修正変化と, 実際に計測された実 ZMP の x 成分 (“act ZMP x”) の時間変化のプロットを図 7.10 に示す. 軌道修正がまだ実行されていない 0[s]-2[s] の期間における最終目標 ZMP (“fin ref ZMP x”) の変化は, 既存の Linear Inverted Pendulum Tracking Controller [55] による修正 $p^d \rightarrow p^{d*}$ を表しており, 2[s] 以降は提案した軌道修正による更なる修正 $p^{d*} \rightarrow p^{d**}$ が与えられている. 最終目標 ZMP と実 ZMP の間のエラーを観察すると, Linear Inverted Pendulum Tracking Controller [55] のみの適用時 (0[s]-2[s] の区間) においては追従が保たれていないエラーが,

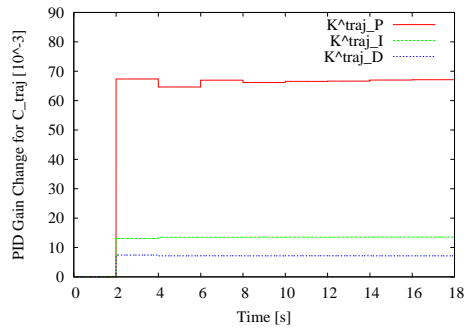
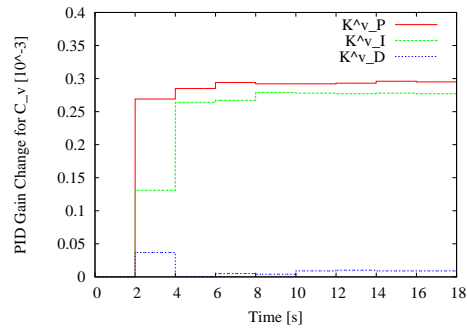
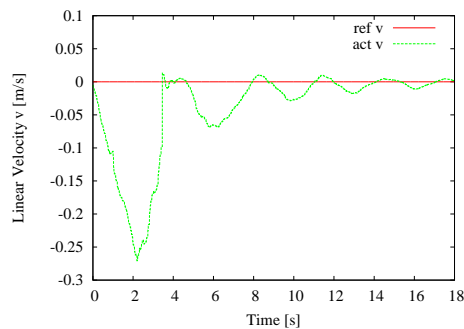
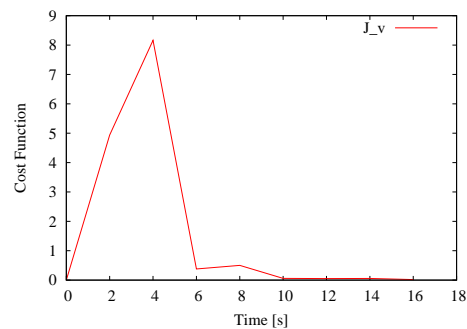
(a1) $t = 0.0[s]$ (a2) $t = 2.0[s]$ (a3) $t = 8.0[s]$ (a4) $t = 18.0[s]$ (b) $K_p^{traj}(n), K_I^{traj}(n), K_D^{traj}(n)$ (c) $K_p^v(n), K_I^v(n), K_D^v(n)$ (d) v^d, v (e) J_v

図 7.9: Experimental Results for Speed Governing Behavior by Life-Sized Humanoid Robot on “Ninebot One S2”

In addition to AutoBalancer [85] and Linear Inverted Pendulum Tracking Controller [55], we introduce the proposed hybrid system of trajectory modification for reference ZMP and gain learning for speed governing behavior control system.

軌道修正の適用後（18[s] 付近）にはエラーを小さく抑えることに成功している．このように，一般的な床面上から搭乗型機器（本実験の例では“Ninebot One S2”）上へと環境が変化した際に，従来の安定化制御手法ではカバーしきれない要因に対しても適応的に軌道修正を行うことの有用性が示された．

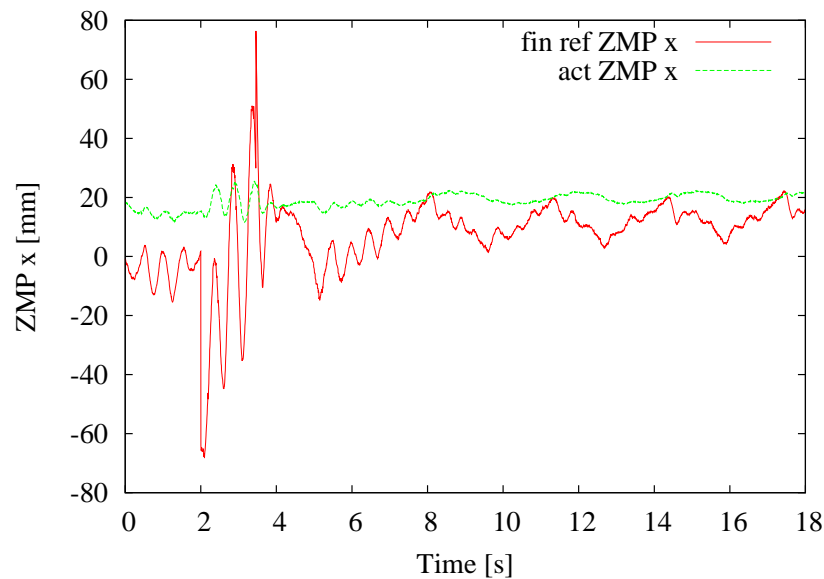


図 7.10: The Final Reference ZMP_x (Red Solid Line) and The Actual Measured ZMP_x (Green Dashed Line) under Speed Governing

学習結果を利用した補助輪のない電動一輪車上でのヒューマノイドの運転行動実現

図 7.9 で示した軌道修正とゲイン学習に対して獲得された PID ゲインの学習結果を利用することで，ヒューマノイドによる電動一輪車“Ninebot One S2”の運転行動を実現する．本実験では“Ninebot One S2”の補助輪を取り外し，図 7.11 に示されるようにヒューマノイド実機を Roll 方向に不安定な“Ninebot One S2”に搭乗させた上で走行実験を行う．図 7.11 のキャプチャでは，一時的に横から人間の手により転倒防止のための補助を与えている．この状態において，図 7.8 のシステム中の並進速度の目標値 v^d に $v^d = 0.3[\text{m/s}]$ のステップ指令を与えることで走行を行う．また，Roll 方向のバランスを制御する用途で，ヒューマノイド実機の Chest フレームの Roll 成分の目標姿勢角 ϕ_C^d に対する目標値整形を式 7.6 のように与

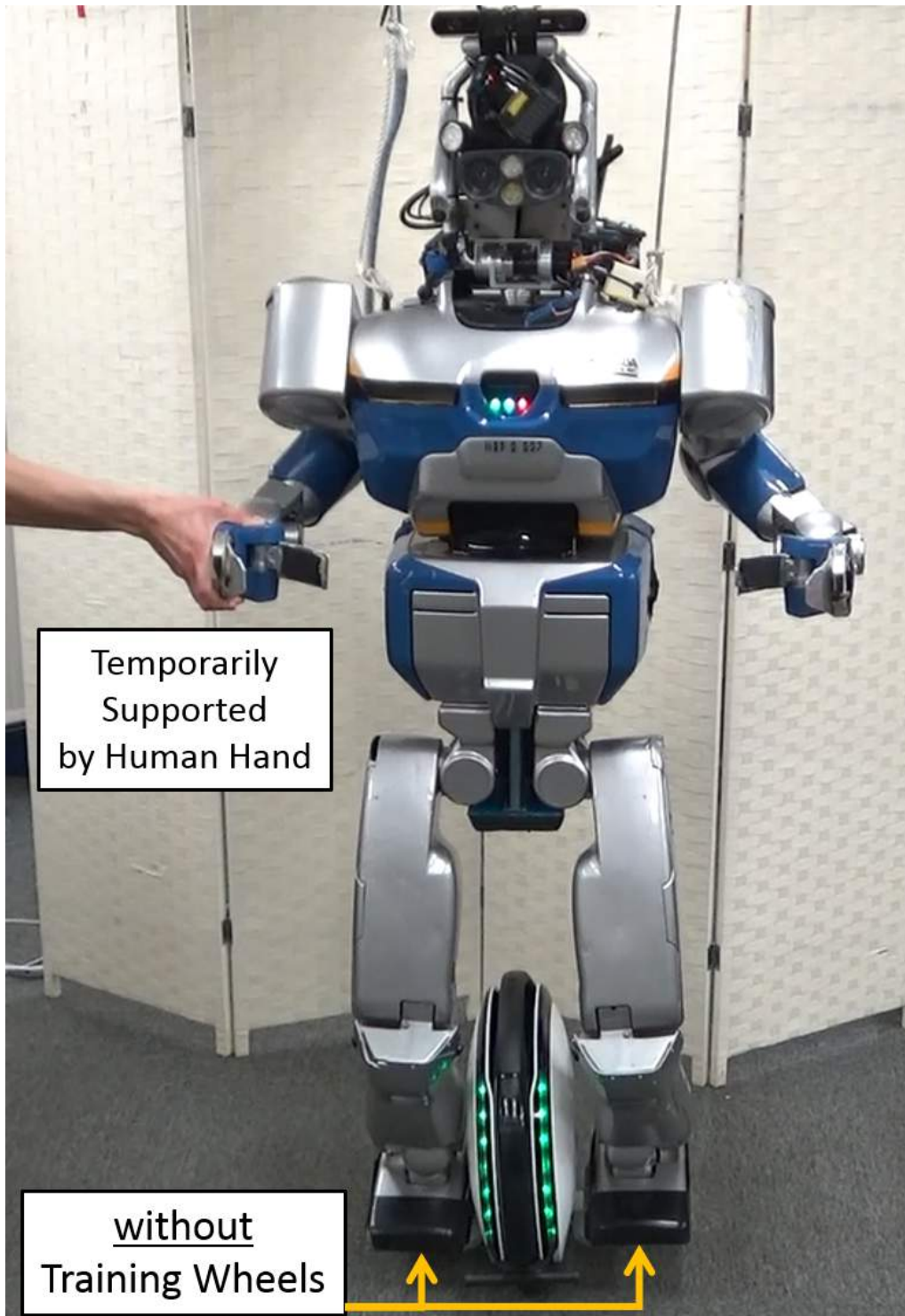


図 7.11: Overview of Life-Sized Humanoid Robot on “Ninebot One S2” without Training Wheels

える。

$$\phi_C^d := \phi_C^d + \alpha \varepsilon_{\text{joy}}, \quad \alpha > 0, \quad \varepsilon_{\text{joy}} \in [-1, 1] \quad (7.6)$$

α は正定数であり, ε_{joy} は人間の Joystick コントローラによる操縦指令から与えられる-1~1の範囲を自在にとる実数である. 式 7.6 の操縦インターフェースに従い目標値 ϕ_C^d が整形される. ϕ_C^d の値に応じて式 4.8 で示した制御則により Chest フレームの Roll 成分の実姿勢角 ϕ_C が制御されることでバランスをリアルタイムに補正する.

運転行動実験のキャプチャを図 7.12 に示す. 図 7.12 -(a1)-(a6) が運転行動中の時系列のキャプチャを表す. 補助輪を外した電動一輪車“Ninebot One S2”の Roll 方向に関しては, 後述の第 8 章で整理する図 8.1 の搭乗型機器の分類において本論文の手法適用が困難となる課題であるが, 手動により目標値を整形する操縦インターフェース (式 7.6) や図 7.12 -(a6) のように人手による介助を割り込むことで走行行動を可能とした. このときの前方への並進速度の目標指令値 v^d (“ref v”) と計測値 v (“act v”) の時間変化を示したプロットが図 7.13 である. 運転行動中の走行挙動に関しては, 図 7.12 -(a1)-(a6) および図 7.13 の応答からも観察できる通り, 目標を満たす速度制御を実現していることが確認できる.

7.4 切替処理による行動実現

7.4.1 搭乗型機器の運転中において割り込み処理型座り直し制御を実行する切替処理実験

第 3 章で述べた搭乗型機器の操作に対する軌道修正において, 4.3 節の搭乗姿勢の安定化を目的とした割り込み処理型座り直し制御 (4.3.2 項) による軌道修正を導入した, オンライン探索指標の切替処理による行動実現を示す. 表 3.3 の搭乗型機器の操作におけるオンライン探索指標ならびに軌道修正のパラメータ対応から, 表 4.3 で示した割り込み処理型座り直し制御におけるパラメータ対応へと, 表 7.5 に示されるような切替処理を実行する.

本行動の定性的な特徴として, 7.3.1 項で示した安定化を図る並行処理実験とは対照的に,



図 7.12: Experimental Snapshot of Driving “Ninebot One S2” without Training Wheels by Life-Sized Humanoid Robot

運転操作を途中で中断させた安定化処理を割り込むことで

運転操作の一時中断

⇓ オンライン探索指標の切替処理

割り込み処理型座り直し制御による搭乗姿勢の安定化

⇓ オンライン探索指標の切替処理

運転操作の再開

に従ったプロセスで安定化行動を実現する。

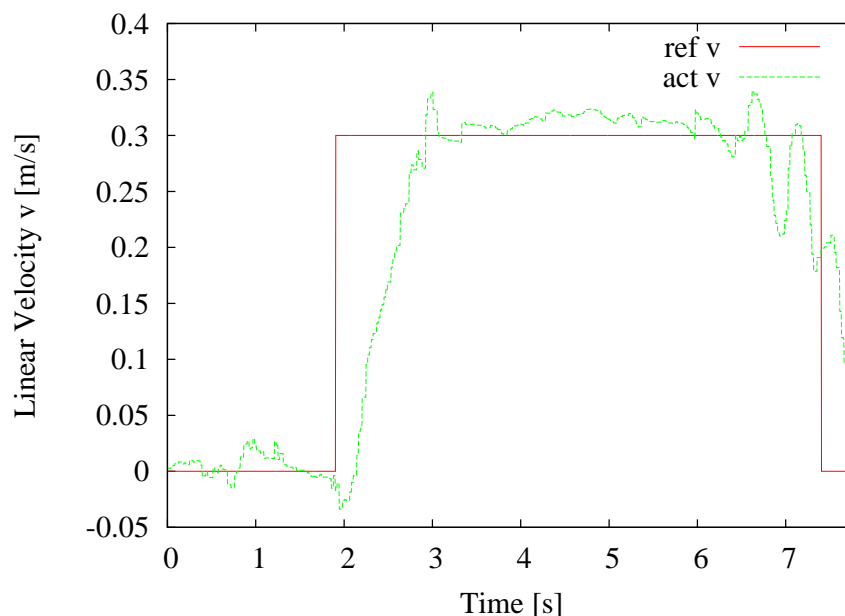


図 7.13: Reference and Response v^d , v for Linear Velocity Control while Driving

The reference velocity v^d (red solid line) is the step input, and the actual measured velocity response v (green dashed line) is controlled by the trajectory modification and the speed governing behavior control based on learned PID gains.

搭乗型機器の運転操作中に、割り込み処理型座り直し制御へと切替処理を実行することで安定化を図る行動実験を図 7.14 に示す。図 7.14 では、左側の (fail-1)-(fail-4) に、切替処理を実行しなかった場合の搭乗型機器の操作のみによる検証実験のキャプチャを表し、右側の (success-1)-(success-4) に、割り込み処理型座り直し制御への切替処理を加えた場合の安定化行動実現を表している。

オンライン探索指標の切替処理を与えない場合（表 7.5 の“搭乗型機器の操作”の 1 処理のみ）の通常の運転操作では、4.3.1 項で詳説した Roll まわりの体幹傾斜が運転中に徐々に蓄積されており（図 7.14 -(fail-1) から図 7.14 -(fail-3)），その影響で図 7.14 -(fail-4) のキャプチャで観察できるようにヒューマノイドは搭乗型機器である三輪車の車体ごと横転してしまっている。

一方で切替処理（表 7.5 で示す“搭乗型機器の操作”から“割り込み処理型座り直し制御”への切替遷移）を加えた場合の行動実験では、ヒューマノイドが運転操作中に体幹の Roll 傾斜を検知し（図 7.14 -(success-2)），それに基づいた割り込み処理型座り直し制御を 4.3.2 項で述べた方策により実行することで傾斜を修正した安定化行動を実現している（図 7.14 -

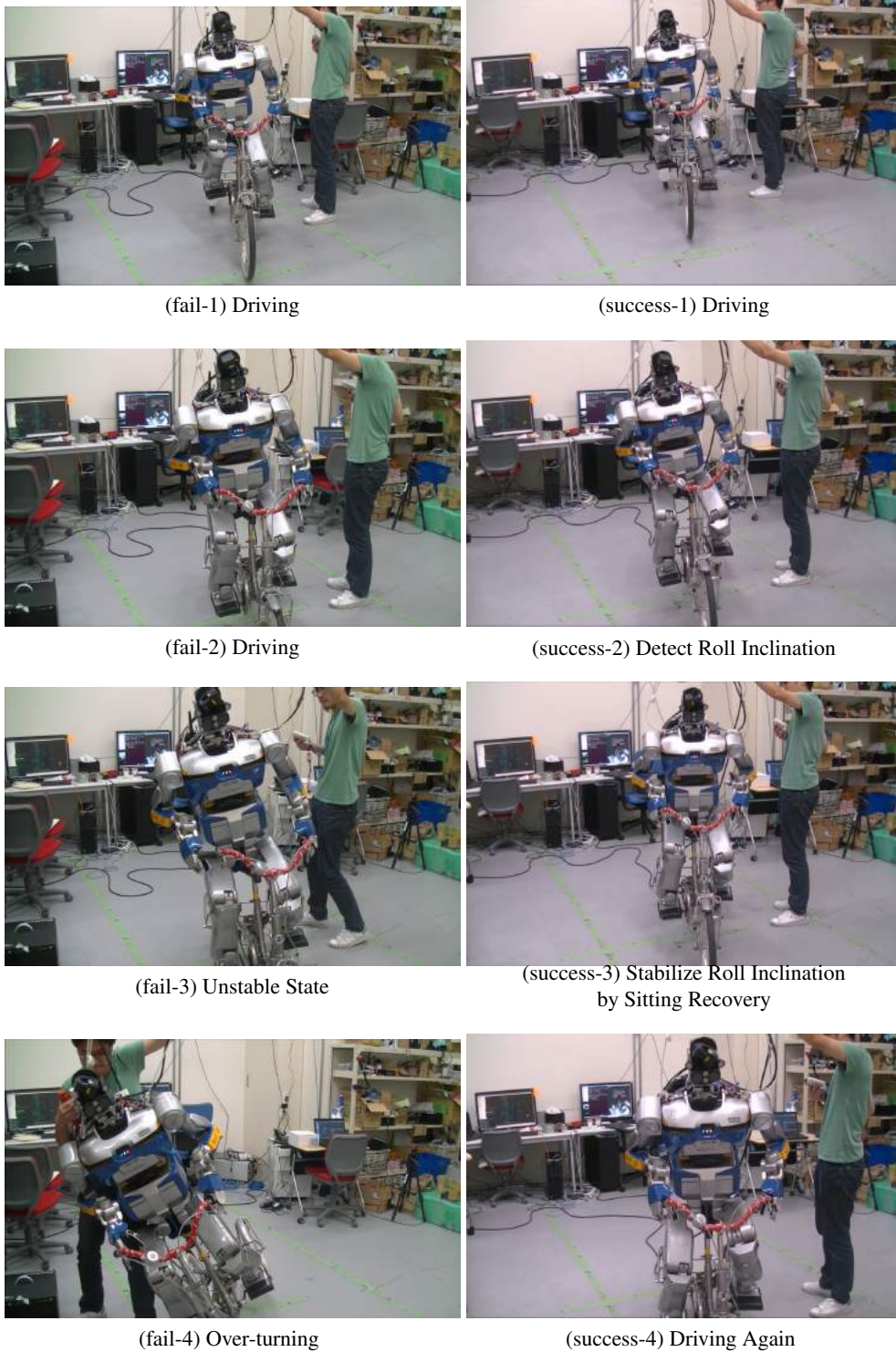


図 7.14: Switching Behavior of Vehicle Manipulation and Sitting Recovery Stabilization (success-1)-(success-4) Compared with Default Vehicle Manipulation (fail-1)-(fail-4)
Proposed sitting recovery motion is executed from (success-2) to (success-3).

表 7.5: 搭乗型機器の操作と割り込み処理型座り直し制御の切替処理において対応する各パラメータ

切替処理	搭乗型機器の操作	割り込み処理型座り直し制御
オンライン探索指標 $f(t; n)$	$\text{sgn}[\rho(t; n), f^{\text{msr}}(t)](f^{\text{msr}}(t) - f^{\text{trg}}(t))$	$\text{sgn}[\rho(t; n), \theta_{\text{roll}}(n)](\theta_{\text{roll}}(n) - 0)$
更新される軌道 $\rho(t; n)$	$\xi^{\text{ref}}(t)$	$y_g(n)$
一般化更新則の重み A	$(M_v s^2 + D_v s + K_v)^{-1}$	$L_{\text{hip} \rightarrow \text{COG}}$

(success-3) . 表 4.5 で整理した割り込み処理型座り直し制御の特徴として挙げられている, ヒューマノイドの Base フレーム (骨盤リンク系) も含めた姿勢修正が実現されている点も, 図 7.14 -(success-2) → 図 7.14 -(success-3) の変化から観察できる. 安定化行動の完了後は, 再び運転操作のフェーズへと戻るオンライン探索指標の切替処理 (表 7.5 で示す“割り込み処理型座り直し制御” から “搭乗型機器の操作” への切替遷移) を実行することで, 図 7.14 -(success-4) に示すように継続的に運転の再開を行っている.

以上のように本行動実験では, 表 7.5 に従った以下の2つのオンライン探索指標の切替処理に基づく軌道修正が達成されている.

- $\text{sgn}[\rho(t; n), f^{\text{msr}}(t)](f^{\text{msr}}(t) - f^{\text{trg}}(t)) \Rightarrow \text{sgn}[\rho(t; n), \theta_{\text{roll}}(n)](\theta_{\text{roll}}(n) - 0)$:

図 7.14 -(success-2) → 図 7.14 -(success-3)

- $\text{sgn}[\rho(t; n), \theta_{\text{roll}}(n)](\theta_{\text{roll}}(n) - 0) \Rightarrow \text{sgn}[\rho(t; n), f^{\text{msr}}(t)](f^{\text{msr}}(t) - f^{\text{trg}}(t))$:

図 7.14 -(success-3) → 図 7.14 -(success-4)

以下では, 図 7.14 -(success-2) → 図 7.14 -(success-3) の期間での割り込み処理として実現されている座り直し動作の効果について詳しく検証する.

座り直し動作による傾斜の修正結果

図 7.14 -(success-2) → 図 7.14 -(success-3) における座り直し動作のキャプチャを図 7.15 の左図に示す. 図 4.5 で示した双脚の能動機能による踏み込みの目標接触力は, 式 4.2 の $F_{\text{hip},z}$

が0となり臀部が浮上して全体重と拮抗するような $f^{trg}(t) = (0 \ 0 \ 320)^T$ [N] の値に設定している。

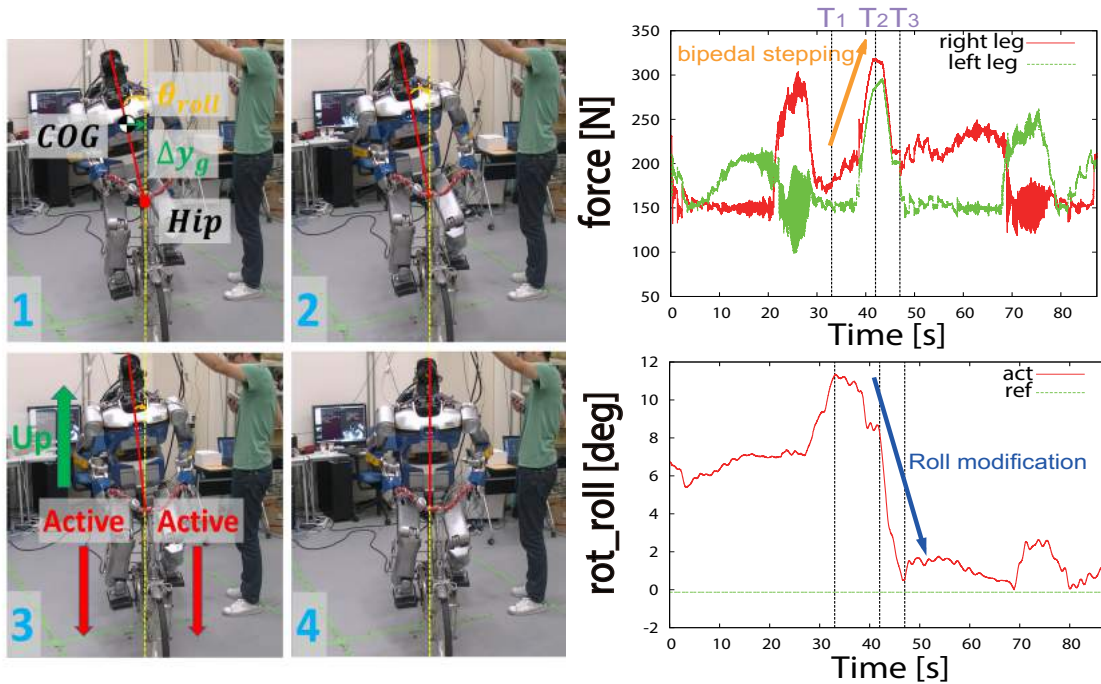


図 7.15: Sitting Recovery Motion by Applying Bipedal Active Functions [52]

Top right plot shows the change of vertical reaction force. Bottom right plot shows the change of roll attitude angle.

図 7.15 の右図の時系列プロット（上:足がペダルから受ける鉛直成分反力，下:体幹のロール傾斜角）に示すように，時刻 T_1 で補正時期 $\theta_c = k\pi$ (k : 整数) の判定かつ姿勢のズレを検知後， $T_1 \sim T_2$ で目標接触力に向けて双脚の能動的な踏み込みをおこない臀部とサドルの接触離脱を実現させる．その後， T_3 まで式 4.3 に従った重心補正を行った．図 7.15 の右下のグラフをみると，ロール傾斜角が T_3 以降で低減されていることがわかる．また，図 7.15 の第 4 スナップショットのように座面ズレが解消されている．先に述べたように座り直し動作後 (T_3 以降) は再びペダリングによる運転操作を再開している．この際，第 3 章で論じたマニピュレータの能動機能と受動機能に基づいた双脚によるペダリング操作を実現していることが，図 7.15 の右上で示した双脚の計測接触力のグラフから判断できる．以上の時系列動作を表 7.6 のタイムテーブルに整理する．

表 7.6: Time Table about Sitting Recovery Motion

Time [s]	Behavior
$0 \sim T_1$	Drive until detecting roll angle threshold and checking crank angle $\theta_c = k\pi$ (k : integer)
$T_1 \sim T_2$	Generate bipedal active standing → Cancel hip contact force
$T_2 \sim T_3$	Move COG of humanoid → Recover sitting
$T_3 \sim$	Drive again based on active and passive pedaling

座り直し動作による転倒・負荷の解消とズレの補正効果

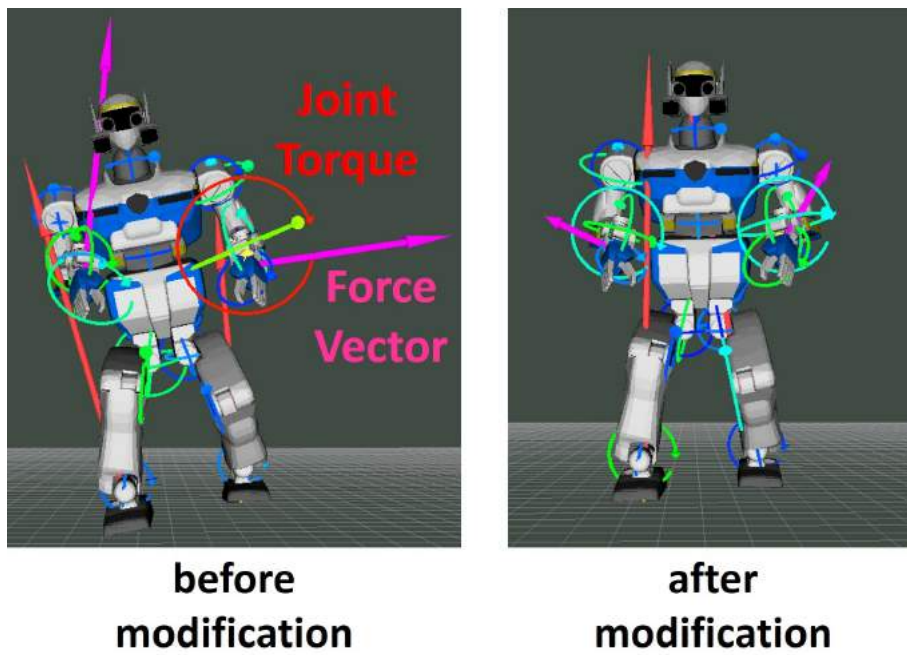
4.3.1 項で議論した座面ズレで提起される問題について、(問題 1)～(問題 3) の課題が、座り直し動作により解消されていることを検証する。図 4.3 の右図でも示した搭乗中のヒューマノイドにおける負荷状態の監視モニタを図 7.16 -(I) に示す。図 7.16 -(I) の左図が座り直し動作を実行する前の傾斜した状態であり、図 7.16 -(I) の右図が座り直し動作達成後の傾斜補正された状態を表す。図 4.3 と同様に図 7.16 -(I) の左図からは(問題 1),(問題 2) が観察される。座り直し動作後の図 7.16 -(I) の右図では、重心位置が中央に寄り、力・関節トルクの大きさ(ノルム)も低減され、負荷が解消されていることが分かる。

次に、三輪車と搭乗したヒューマノイドの相対関係について、

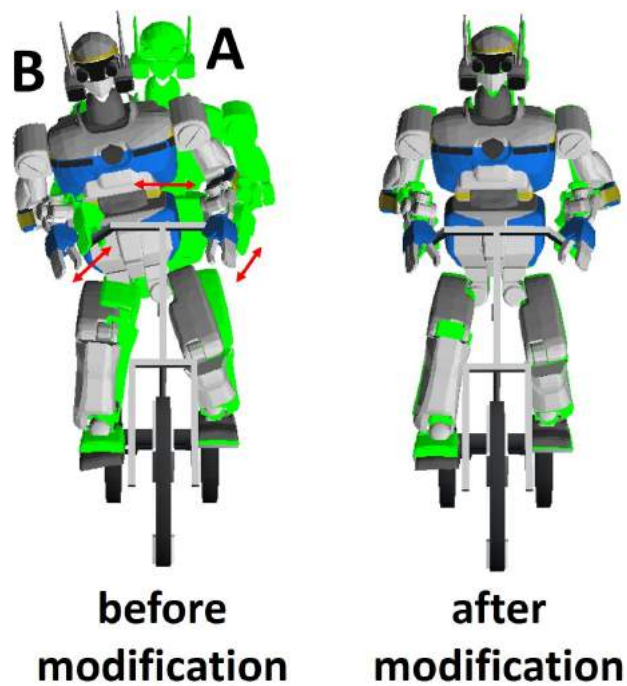
A 搭乗したヒューマノイドの腰部位置姿勢、関節角度の目標値を反映したモデル

B 搭乗したヒューマノイド実機の体幹姿勢推定に基づく姿勢情報、および関節エンコーダによる各関節角度を反映した実環境ロボットモデル

の 2 つの差分を比較評価した図が図 7.16 -(II) である。図 7.16 -(II) の左図が座り直し動作前の傾斜した姿勢、図 7.16 -(II) の右図が座り直し動作後の状態である。ヒューマノイド実機では、体幹姿勢推定を利用した姿勢、およびエンコーダより計測された関節角度の情報が取得可能であるが、位置情報については計測が困難である。ここではヒューマノイドの手先足先位置のズレの傾向を把握することを目的とし、モデル A とモデル B の両足支持点(各足の圧力中心の中心)の位置を図 7.16 -(II) では一致させている。この条件に基づき、両足支持点を位置相対とした場合の手先および腰部のズレを観察する。



(I) Monitoring Viewer about Force Vector and Joint Torque



(II) Difference between Reference Model A and Real Robot Model B

図 7.16: Comparison about Loads and Position Errors before and after Sitting Recovery Motion [52]

補正前の図 7.16 -(II) の左図においては、目標値を表しているモデル A の手先の把持箇所・腰部位置姿勢と、実環境ロボットモデル B の手先・腰部のズレが顕著である。このような接触ズレを保った状態でハンドルやクランクへの操作行動を実行すると、目標操作量（計画軌道）の誤差伝搬や接触箇所の脱落へとつながりうる。これが 4.3.1 項で述べた（問題 3）に該当する。座り直し動作を実行することで、ヒューマノイド関節の角度状態のみならず、制御の難しい Base フレーム（腰部の骨盤リンク系）の搭乗位置・姿勢に関しても修正が可能となり、図 7.16 -(II) の右図のように A と B が十分近くなったことが確認できる。

7.4.2 搭乗型機器への乗り込み動作と搭乗後の调速行動をシーケンシャルに実行する切替処理実験

これまでに論じてきた行動実現では、オンライン探索指標における分類の内、特にオンライン制御指標を利用した軌道修正の並行処理と切替処理について述べた。本項では、オンライン探索指標の中でもオンライン学習指標に焦点を当てたゲイン学習の切替処理について述べる。

オンライン探索指標（オンライン学習指標）を利用したゲイン学習法として本論文では、第 5 章で提案した搭乗する環境に適応した足部トルクダンピング制御系に対するゲイン学習と、第 6 章で提案した搭乗した環境の運動に適応した调速行動制御系に対するゲイン学習について、それぞれ独立したオンライン学習制御系を設計した。本項において目的とする行動実現では、これら 2 つの独立したオンライン学習制御系をシーケンシャルに切り替えるオンライン探索指標の切替処理法を提案し、切替処理を利用することでヒューマノイドの搭乗型機器への乗り込み動作と乗り込んだ後の運動に適応する调速行動を継続した複合タスクを実行する。第 5 章では、搭乗型機器への乗り込み動作における即時的なゲイン学習を実現する SGD-based open-loop Learning Control (SLC) の手法を提案しており、SLC におけるオンライン探索指標とゲイン学習の一般化更新則とのパラメータ対応は表 5.1 で整理した。第 6 章では、搭乗後を前提とした搭乗環境の運動に適応するゲイン学習法として Mini-batch-based closed-loop Learning Control (MLC) の手法を提案しており、MLC におけるオンライン探索指標とゲイン学習の一般化更新則とのパラメータ対応は表 6.4 で整理した。

以上に整理した 2 つのオンライン学習制御系をシーケンシャルに結合した切替処理においては、表 7.7 に示されるような 2 つの処理のオンライン探索指標を切り替える必要がある。

表 7.7: SLC に基づく乗り込み動作と MLC に基づく調速行動の切替処理において対応する各パラメータ

切替処理	SLC に基づく 搭乗型機器への乗り込み動作	MLC に基づく 搭乗後の調速行動
オンライン 探索指標 $f(t; n)$	$\left(\frac{\partial J(e_T(t), \delta(t))}{\partial \rho(t; n)}\right)^T$	$\left(\frac{\partial J(v, \omega)}{\partial \rho(t; n)}\right)^T$
更新される ゲイン $\rho(t; n)$	D	ρ_n
一般化更新 則の重み A	η	$\text{diag}(\eta_P, \eta_I, \eta_D)$

Sequential Online Learning Control (SOLC)

搭乗型機器に対してヒューマノイドがシーケンシャルな適応行動を実現することを目的とした、表 7.7 に示す切替処理に準じた 2 つのオンライン学習制御系の切替方を提案する。図 5.3 で提案した SLC による学習制御系のアプローチと、図 6.10 で提案した MLC による学習制御系のアプローチをシーケンシャルに処理することで、図 7.17 で示されるような切替変化可能なオンライン学習制御系が構築できる。本研究では、この切替変化型のオンライン学習制御法を“Sequential Online Learning Control (SOLC)”と定義する。

SOLC は、MLC と SLC の 2 つの学習制御系のカスケード結合で構成されており、その結合点となる適応行動の変化点はヒューマノイドが搭乗型機器に移った直後の状態に相当し、本行動実験ではヒューマノイドの両足が搭乗型機器である平行二輪電動スクーター上に着地した時点とする。着地の状況把握は、図 5.5 で提案したように足部搭載の 6 軸力センサから足裏反力の値を検出し判定を行う。したがって、変化点の前のフェーズが乗り込み動作のフェーズ、変化点の後のフェーズが調速行動のフェーズとして切り分けられる。

以上の変節構造を有した搭乗型機器に対するヒューマノイドの適応行動生成のフローを図 7.18 に示す。図 7.18 において SOLC が適用される範囲は、乗り込み動作生成から乗り込み動作完了後の調速行動に基づく停止操作（並進速度および旋回角速度を零へと漸近させる操作）の行動フェーズまでとし、第 5 章で提案した SLC によるダンピングゲインのオンラインチューニングと、第 6 章で提案した MLC による調速行動制御系の PID ゲインのオンラインチューニングの 2 つの処理のオンライン探索指標を切り替えることでシーケンシャル

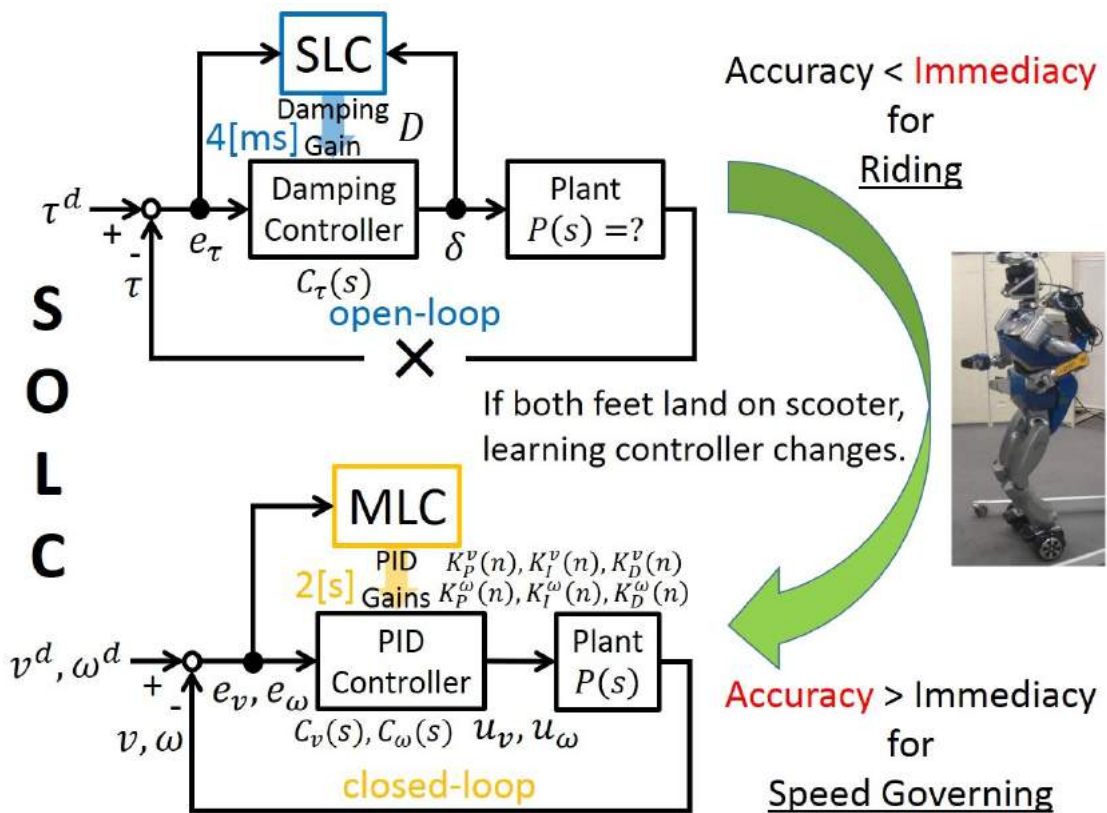


図 7.17: SOLC Approach

にゲイン学習を実行する。

SOLC に基づく搭乗型機器への乗り込み動作と调速行動のシーケンシャル行動実現

SLC と MLC を結合した SOLC による継続的ゲイン学習を利用した、等身大ヒューマノイドロボット HRP2-JSK [110] による平行二輪電動スクーターへの乗り込み動作と调速行動の継続行動実験を示す。

図 7.19 は、平行二輪電動スクーター前方の床面上にヒューマノイドが直立した状態 (Initial State) から、乗り込み動作を経て平行二輪電動スクーター上で调速行動を達成するまでの実験キャプチャを表し、図 7.20 は、そのときの学習対象となる各ゲインおよび評価関数の時間変化を示している。実験キャプチャ (図 7.19) に示すように、Initial State ($t = 0[s]$) か

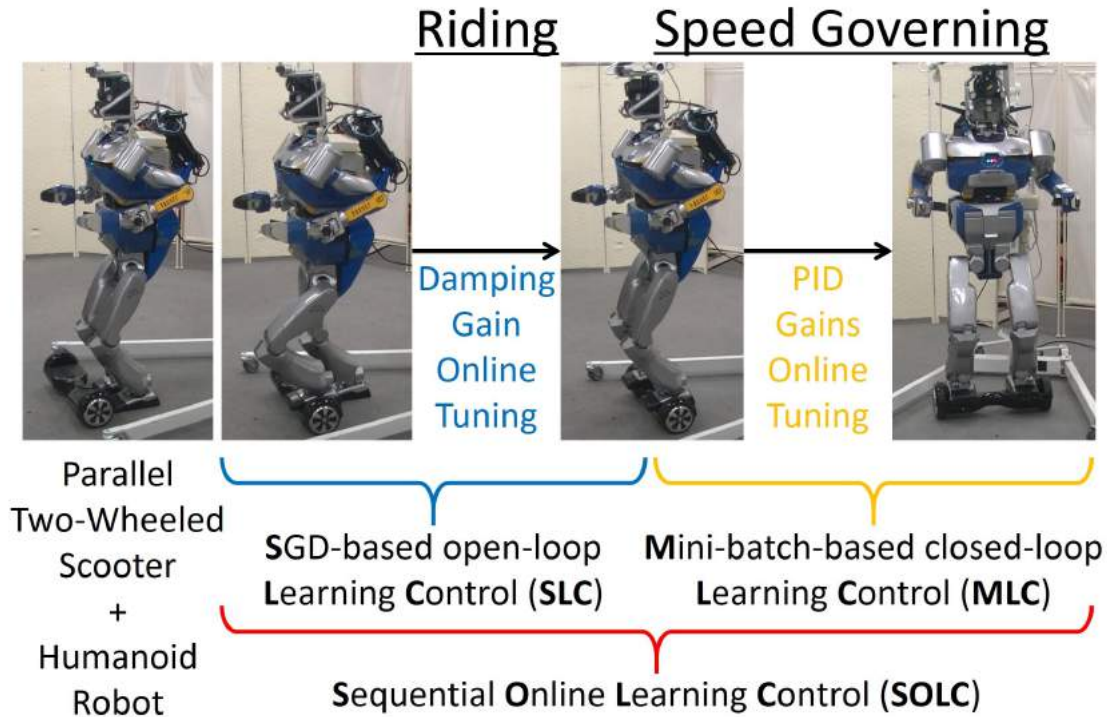


図 7.18: Riding and Speed Governing for Parallel Two-Wheeled Scooter Based on SOLC by Humanoid Robot
SOLC is composed of SLC and MLC.

ら調速行動による停止操作が完了するまでは全体で 79.5[s] かかり， $t = 79.5[s]$ で式 6.32 の終了条件を満たした．SOLC は $t = 4[s]$ から開始され， $t = 5.5[s]$ を SLC と MLC の変化点として，乗り込み動作のフェーズ（Riding Phase）と調速行動のフェーズ（Speed Governing Phase）の時間範囲は以下となった．

$$\left\{ \begin{array}{ll} \text{Riding Phase by SLC :} & 4 \leq t \leq 5.5 \\ \text{Speed Governing Phase by MLC :} & 5.5 \leq t \leq 79.5 \end{array} \right.$$

図 7.20 -(a) は，Riding Phase における左足 Pitch の足部トルクダンピング制御系のダンピングゲインの時間変化のプロットを表している．図 7.20 -(b) および図 7.20 -(c) は，Speed Governing Phase における v と ω に対する調速行動制御系の PID ゲインの時間変化のプロットを表す．SLC における学習率等の設定パラメータは表 5.2 の値を適用し，MLC における v と ω に対する学習率はそれぞれ表 6.6 の設定値に従う．

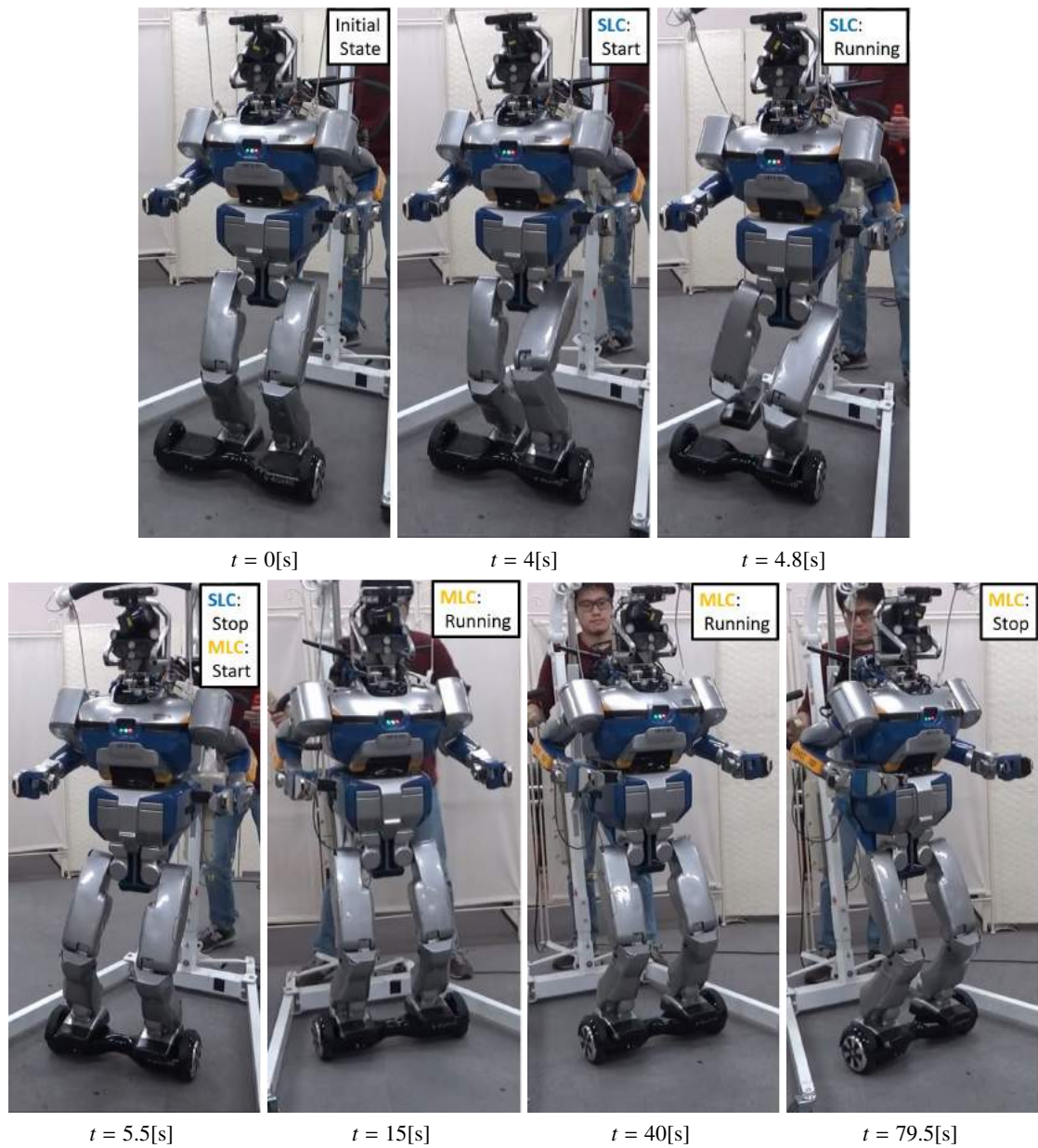


図 7.19: Snapshot for Sequential Experiment of Riding and Speed Governing
 SLC is running from $t = 4[s]$ to $t = 5.5[s]$. MLC is running from $t = 5.5[s]$ to $t = 79.5[s]$.

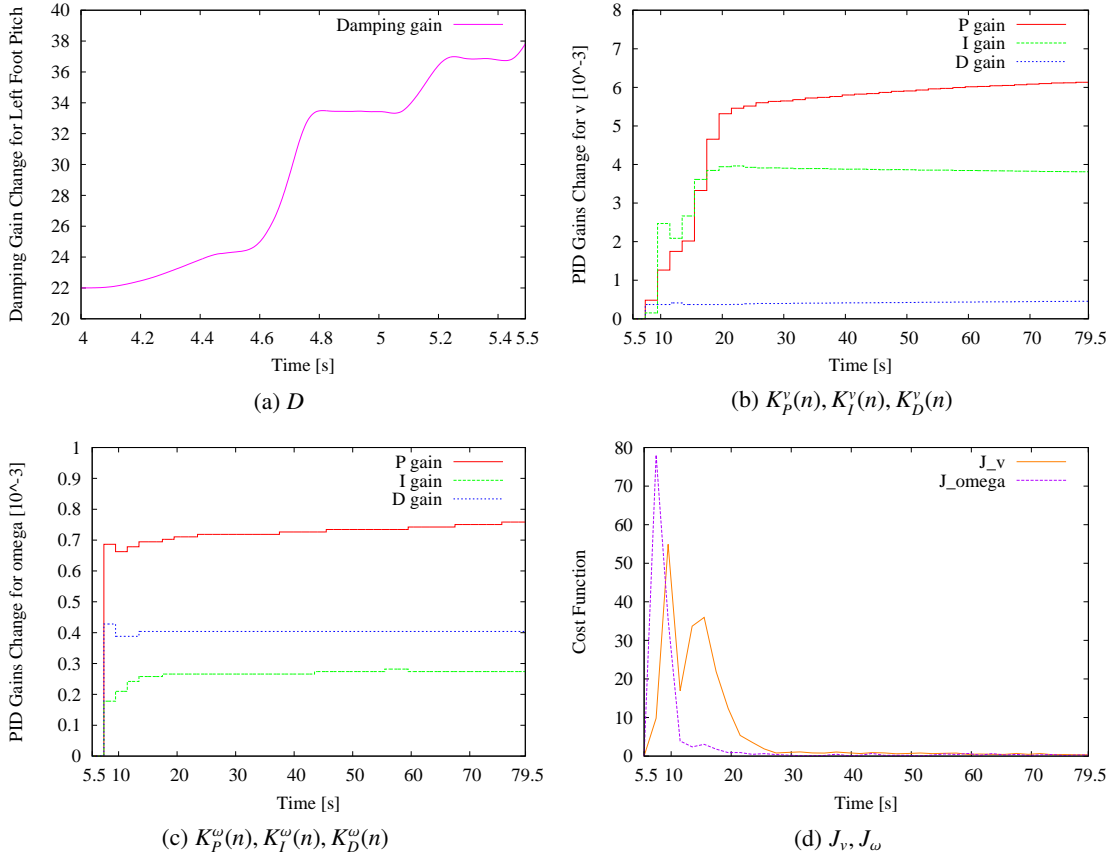


図 7.20: Change of Controller Gains and Cost Functions for Sequential Experiment of Riding and Speed Governing

Each change represents (a) damping gain, (b),(c) PID gains for v, ω , and (d) MLC cost functions for v, ω . SLC is running in 4[ms] cycle from $t = 4$ [s] to $t = 5.5$ [s]. MLC is running in 2[s] cycle from $t = 5.5$ [s] to $t = 79.5$ [s].

ダンピングゲインの学習結果に関しては、図 5.8 で得られた実験結果と同様に増加傾向を呈しており、各 PID ゲインの学習結果に関しては、初期値を 0 として MLC を開始させた後、ゲインの値が収束に向かっていくことが観察できる。また、図 7.20 -(b) の学習結果より、 v に対する调速行動制御では D 制御に比べて P 制御と I 制御が支配的となる結果が得られた。このときの MLC の評価関数 J_v, J_ω の時間変化が図 7.20 -(d) である。MLC の開始初期では J_v, J_ω ともに評価関数が活性化しており、MLC が進行するにつれて 0 に向かって収束している様子が確認できる。

図 7.21 は、初期時刻 $t = 0$ [s] から停止操作の完了時刻 $t = 79.5$ [s] までの调速行動制御系の目標値 $v^d = 0, \omega^d = 0$ と制御量 v, ω の時間変化を表す。図 7.21 -(i) と図 7.21 -(ii) の結果から

は、SOLCの適用により最終的に目標値 $v^d = 0, \omega^d = 0$ への収束（停止操作）に成功していることが観察できる。

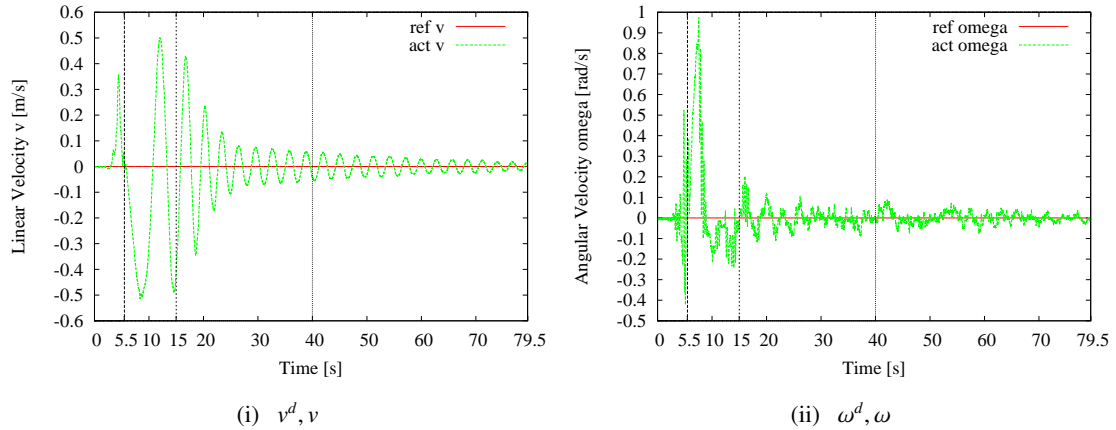


図 7.21: The References $v^d = 0, \omega^d = 0$ and The Actual Measured v, ω

7.5 本章のまとめ

本章では、ヒューマノイドにおける搭乗型機器の運転操作と適応行動で課題となる、“機器の操作”・“搭乗姿勢の安定化”・“外界環境の認識”・“搭乗する環境への適応”・“搭乗した環境の運動への適応”，が複合的に組み合わさる統合課題に対して、オンライン探索指標の並行処理と切替処理の2つの処理方策を提案し、実装したそれぞれの処理に基づいた行動実現を示した。

オンライン探索指標を並行処理として扱うことで、同じ制御ループ内で共存したタスク実現が可能であり、本章では運転操作の要となる“機器の操作”に対して“搭乗姿勢の安定化”と“外界環境の認識”を同時実行させた行動を実現した。ヒューマノイドがもつ大きな特徴として、多様な物理量を計測・把握するための複合的なセンシングシステムが挙げられ、本章の行動実験で示したオンライン探索指標の並行処理においては、

- “機器の操作”で働くヒューマノイドと搭乗型機器の間に作用する力覚情報
- “搭乗姿勢の安定化”に要求されるヒューマノイド自らの慣性情報
- “外界環境の認識”を目的とした視覚ナビゲーション情報

を同じ処理内で扱うことにより、複数の探索変数を共通のシステム上で利用することを可能とした。さらに、軌道修正とゲイン学習の2つの異なる概念が共存可能となるような並行処理システムを提案し、これまで人間が行動タスクの観察に応じて軌道修正を適用すべきか、ゲイン学習を適用すべきかを思考・決定していたプロセスから自律させた適応行動実現を示した。

一方、オンライン探索指標を切替処理として扱うことにより、別々の行動戦略が互いに干渉しない切替遷移に基づいたシステムを構築することが可能となり、状況判断に応じて適応的に行動戦略を変節することができる。本章で示したオンライン探索指標の切替処理に基づくゲイン学習方策では、

- “搭乗する環境への適応”を節義とした足部トルクダンピング制御系に対する学習戦略
- “搭乗した環境の運動への適応”を節義とした調速行動制御系に対する学習戦略

の2つの処理を切り替えた変節により、搭乗型機器に対する乗り込み行動と調速行動の一貫した継続行動実現を可能としている。

第8章

結論

8.1 本研究の総括

本研究では、ヒューマノイドにおける搭乗型機器の運転操作と適応行動を実現するための課題として、“機器の操作”・“搭乗姿勢の安定化”・“外界環境の認識”・“搭乗する環境への適応”・“搭乗した環境の運動への適応”の5つの課題を提起し、これらの課題を共通に解くためのオンライン探索指標に基づく軌道修正とゲイン学習の方策および行動戦略を提案した。本論文の各章の概要について以下にまとめる。

第1章「序論」では、ヒューマノイドにおける従来の行動タスクの実現を背景として、本研究が目的とする搭乗型機器への運転操作と適応行動において解くべき課題について述べた。また、それらの課題に対する解として、本研究の主軸となるオンライン探索指標に基づいた軌道修正とゲイン学習の考え方を導入した。

第2章「オンライン探索指標に基づく搭乗型機器の運転操作と適応行動」では、搭乗型機器の運転操作と適応行動をヒューマノイドにより実現させるためのオンライン探索指標の概念を定義し、その適用戦略について軌道修正とゲイン学習の方策によるアプローチを述べた。軌道修正とゲイン学習の考え方を共通に理解するために、一般性を考慮した「軌道とゲインの一般化更新則」を提案した。

第3章「搭乗型機器の操作に対する軌道修正」では、搭乗型機器の運転操作を目的としたヒューマノイドによる“機器の操作”に対する軌道修正戦略について述べた。“機器の操作”ではヒューマノイドにおける搭乗型機器へのマニピュレーション制御法が重要であり、本研究においてはヒューマノイド身体と搭乗型機器で閉じた閉ループ系の構造に着目した能動機能と受動機能のマニピュレータ軌道の制御法について探索指標の観点から提案した。提案したマニピュレーションの軌道戦略に基づき、ヒューマノイドによる双腕・双脚を協調させた三輪車操作行動を実現した。

第4章「運転操作における搭乗姿勢と外界環境に対する軌道修正」では、第3章で実現した運転操作行動に付随する課題として、“搭乗姿勢の安定化”と“外界環境の認識”についてそれぞれ議論した。両者の課題は、どの系からみた軌道修正戦略かを考えることで共通に理解することができる。4.3節では搭乗姿勢の安定化に対する軌道修正法を提案し、4.4節では外界環境の認識に対する軌道修正法を提案した。

第5章「搭乗する環境に適応したゲイン学習」では、ヒューマノイドが属する系の環境変化に対応した“搭乗する環境への適応”に対するゲイン学習法について述べた。“搭乗する環

境への適応”が要求される行動課題として、ヒューマノイドによる搭乗型機器への乗り込み動作を目的とした。即時的な乗り込み動作に対して迅速にゲイン学習が可能な SGD-based open-loop Learning Control (SLC) によるオンライン学習制御法を提案し、ヒューマノイドによる平行二輪電動スクーターへの乗り込み動作を実現した。

第6章「搭乗した環境の運動に適応したゲイン学習」では、ヒューマノイドが搭乗型機器上の環境に属した際の“搭乗した環境の運動への適応”に対するゲイン学習法について述べた。本研究では、搭乗した環境の運動に関して、搭乗型機器上のヒューマノイドが把握する並進速度と旋回角速度に対する調速行動制御系を構成し、調速行動制御系におけるPIDゲインを学習する Mini-batch-based closed-loop Learning Control (MLC) によるオンライン学習制御法を提案した。提案したゲイン学習法に基づきヒューマノイドによる平行二輪電動スクーター上での調速行動を実現した。

第7章「オンライン探索指標の並行処理と切替処理による行動実現」では、第3章～第6章でそれぞれ提案した軌道修正とゲイン学習に基づく行動戦略を複合的に実現させるオンライン探索指標に対する並行処理と切替処理について述べた上で、それらの処理実行による統合的な行動実現を示した。

8.2 本研究の成果

8.2.1 本論文の手法により解かれる搭乗型機器の課題

実世界に存在する搭乗型機器を次のような軸で分類するとき、本論文で提案したヒューマノイドにおける適用手法により共通的に解くことが可能となる領域を図8.1でハイライトとして示す。

安定系搭乗型機器：

搭乗者（人間やヒューマノイド等）が搭乗しない状態において、機器自体が物理的に倒れずに直立することのできる搭乗型機器

不安定系搭乗型機器：

搭乗者（人間やヒューマノイド等）が搭乗しない状態において、物理的支持（スタンド等）のない環境で転倒が生じ姿勢維持が困難な搭乗型機器

能動駆動系搭乗型機器 :

機器に原動機類（電動機やエンジン等）が備わっており，機器自体が能動的に駆動する搭乗型機器

受動駆動系搭乗型機器 :

機器に原動機類（電動機やエンジン等）が備わっておらず，搭乗者（人間やヒューマノイド等）が作用する力学的エネルギーにより受動的に駆動される搭乗型機器

上記の軸に基づき，それぞれの搭乗型機器は 4 つの象限 (I)~(IV) に分類される．本論文の網羅するハイライトの領域が実現可能とする方法論を以下にまとめる．

○ 象限 (I) に対する本論文のアプローチ :

搭乗型機器が有する能動駆動系に対して，ヒューマノイドはその機器の運動の特性を理解し適応する行動が要求される．本論文は，ヒューマノイドが作用する目標 ZMP の軌道修正ならびに調速行動制御系のゲイン学習をオンライン学習制御系 Mini-batch-based closed-loop Learning Control (MLC) を軸に coupling させることで，象限 (I) に属する搭乗型機器への適応行動を可能とした．

○ 象限 (II) に対する本論文のアプローチ :

搭乗型機器が能動駆動により安定化を図るプロセス（安定化制御や安定化機構等）とヒューマノイド自身が安定化を図るプロセスの 2 つが共存した状態での協調的な行動が要求される．本論文は，ヒューマノイドの属する環境が一般的な床面から搭乗型機器上へと遷移する過程において，適応的に安定化制御系のゲインを調整するためのオンライン学習制御系 SGD-based open-loop Learning Control (SLC) を提案した．また，搭乗後の環境においては搭乗型機器がなす安定化のための駆動系とヒューマノイドが実行する安定化制御系が共存可能な調速行動制御系のシステムを構成し，制御系に対するゲイン学習を Mini-batch-based closed-loop Learning Control (MLC) により可能とした．

○ 象限 (III) に対する本論文のアプローチ :

本論文では，象限 (III) に属する搭乗型機器特有の課題を提起し，外部から手動で指令する目標値整形や人間による介助を割り込むことで，ゲインの学習結果を利用した速度指令ベースの制御に基づいて運転行動における部分的な課題を解いた．一方で，不

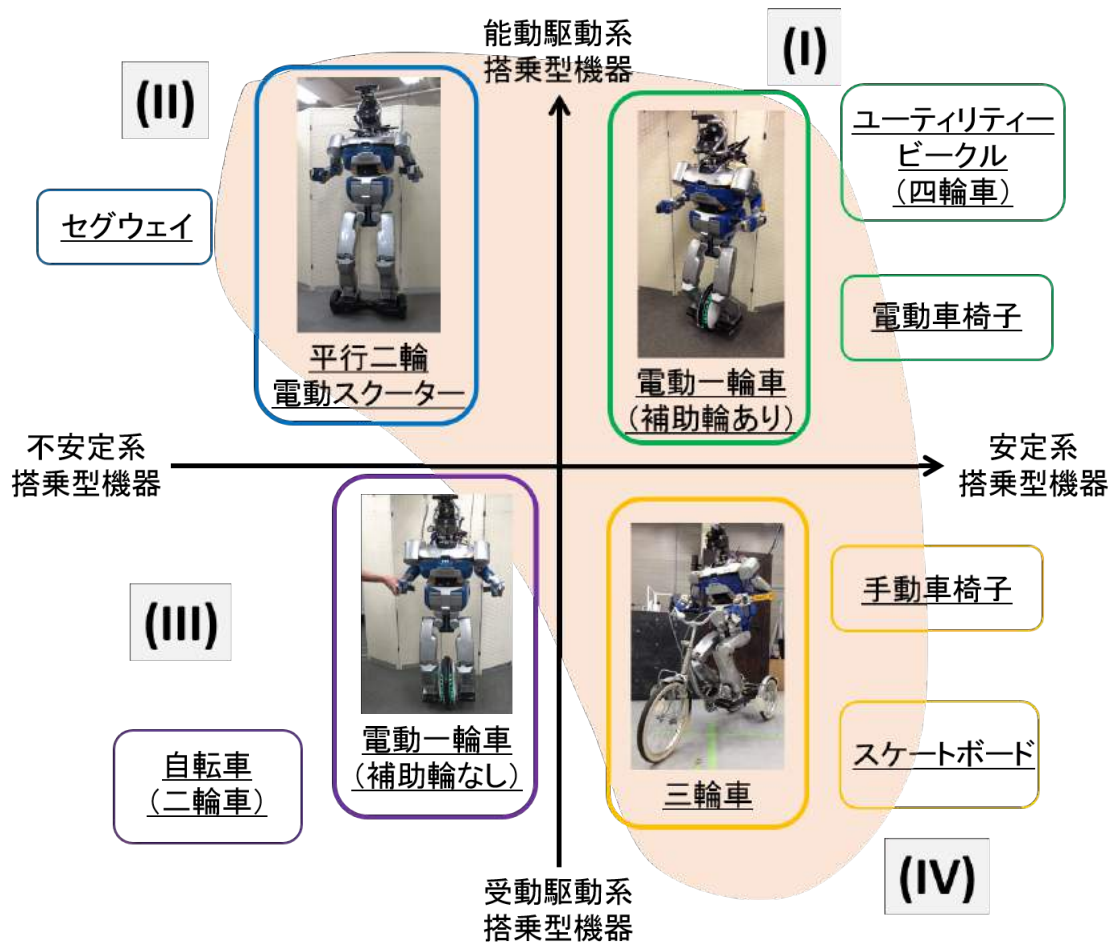


図 8.1: 実世界における搭乗型機器の分類と本論文が解くことを可能とする領域
 安定系搭乗型機器, 不安定系搭乗型機器, 能動駆動系搭乗型機器, 受動駆動系搭乗型機器, を軸にして搭乗型機器は, 象限 (I), 象限 (II), 象限 (III), 象限 (IV) のそれぞれに分類される. 座標平面内においてハイライトで示した領域が, 本論文の提案する方法論によりカバーされ得る搭乗型機器の範囲を表す.

安定系かつ受動駆動系の性質をもつ搭乗型機器においては、搭乗者であるヒューマノイド自身の即応的なセンシング能力・制御処理能力が重要であり、どんな状態量をセンシングし、どのようなパラメータを自律的にチューニングさせるか等の選択的な制御・学習の課題が存在する。

○ 象限 (IV) に対する本論文のアプローチ：

搭乗型機器が有する受動駆動系に対して、ヒューマノイドは機器に対して力学的な作用を適切に供給する行動が要求される。本論文は、ヒューマノイド身体と搭乗型機器の間で発生する操作行動における内力に注目し、この因子を緩和させるように適切に機器を操作する軌道修正法を提案した。また、搭乗型機器自体は安定系であるが、搭乗者であるヒューマノイドは搭乗型機器上で外乱に応じながら適切にバランスを補正することが求められており、この課題に対して搭乗姿勢の安定化を実行する軌道修正法を提案した。最後に、ヒューマノイドと搭乗型機器の閉じた系に限らず、系外の外界環境に対する認識処理に基づいた軌道修正戦略についても提案した。

8.2.2 本研究により得られる波及効果

本研究「ヒューマノイドにおける搭乗型機器の運転操作と適応行動」から、

1. ヒューマノイドの機動性が従来の二足歩行形態から変化することへの対応
2. 事前（オフライン）の設計・計画により対処困難な課題に対するオンライン処理に基づく解法
3. 軌道修正とゲイン学習の2つの異なる方法論の共通化と共存

の波及効果が期待されうると結論付けられる。それぞれの得られる成果について以下に論じる。

1. ヒューマノイドの機動性が従来の二足歩行形態から変化することへの対応
汎用的なヒューマノイドの多くは二足歩行に基づく移動形態を有しているが、二足歩行特有の不安定性や目的地に至るまでの時間・コストに関して考慮すると、必ずしも常に二足歩行形態であることが最適解とは限らない。与えられた状況・場面・環境等を制約条件として目的関数を最適化させるような、より良い移動形態への変節が望ま

れる。本研究により実現された運転操作と適応行動では、不安定性を解消した三輪車操作や二足で直立したヒューマノイドが平行二輪電動スクーターに乗り込み、その移動形態を獲得・学習していく適応的な行動をはじめ、今後のヒューマノイドの移動形態に対する変節に貢献し得る成果が得られている。

2. 事前（オフライン）の設計・計画により対処困難な課題に対するオンライン処理に基づく解法

ヒューマノイドによる行動実現では、外乱の存在や環境の変化がつきものであり、人間が設定もしくは事前獲得・学習したパラメータそのままの適用では対応できない問題が多々ある。本研究においては、ヒューマノイドの大きな特徴の一つである複合的なセンシング能力を活用させたオンライン探索指標に基づくオンラインでの行動処理系を提案した。外乱や環境変化などの要因は、ヒューマノイド自身もつセンサをベースに状態観測・検知・認識が可能であり、これらの情報を適切にフィードバックし行動に移すことで適応的な課題解決が得られる。

3. 軌道修正とゲイン学習の2つの異なる方法論の共通化と共存

ヒューマノイドにおける搭乗型機器の運転操作と適応行動を実現させるために、本研究では軌道修正とゲイン学習の2つの方策・戦略を論じた。運転操作に対しては軌道修正を、適応行動に対してはゲイン学習を、それぞれ分割させることにより適用しているが、本研究では、共通軸となるオンライン探索指標の考え方を基に両者の方法論を一般化させた「軌道とゲインの一般化更新則」（式 2.13）として定義・提案している。これにより、制御の分野における軌道修正の考えと学習の分野におけるゲイン学習の考えを共通に理解することができ、共有したアプローチをとることが可能となった。さらに、これら2つの方策は共存が可能であり、同一のシステム内において複合的に処理することも可能となる。

8.3 本研究の結論

本研究では、ヒューマノイドが搭乗型機器の運転操作および適応行動を実現するために解くべき課題となる、“機器の操作”・“搭乗姿勢の安定化”・“外界環境の認識”・“搭乗する環境への適応”・“搭乗した環境の運動への適応”に対して、共通軸となるオンライン探索指標のアプローチに基づく軌道修正とゲイン学習の行動戦略を提案した。また、軌道修正とゲイン

学習における複合的かつ共存可能な処理方法として、並行処理と切替処理の手法を提案し、ヒューマノイドの統合的な行動実現を示した。

“機器の操作”においては、ヒューマノイドの協調マニピュレーションの閉ループ系構造に着目した能動機能と受動機能のマニピュレータの軌道修正法により搭乗型機器との間に介在する操作行動中の接触内力の解消を実現した。

“搭乗姿勢の安定化”においては、搭乗型機器上で運転操作を行うヒューマノイドの搭乗姿勢の不安定要因に対して、割り込み処理型座り直し制御と実時間並行処理型姿勢制御の2つの安定化処理法を提案し、搭乗姿勢の重心位置・体幹の目標姿勢角に対する軌道修正法に基づき安定化行動を実現した。

“外界環境の認識”においては、ヒューマノイドと搭乗型機器以外の系外に存在する外乱要因を大型障害環境と小型障害環境の2つに分けた上で、視覚や力覚に基づいたナビゲーション情報から搭乗型機器への目標操作量の軌道修正を実行する認識行動システムを提案し、外界環境に対して適応的な運転操作行動を実現した。

“搭乗する環境への適応”においては、ヒューマノイドの属する移動形態が汎用的な二足歩行形態から搭乗型機器上での運動形態へと遷移する際に、安定化のための制御系に対して適応的にゲインチューニングを実行する SGD-based open-loop Learning Control (SLC) に基づくゲイン学習法を提案し、搭乗型機器への乗り込み行動を実現した。

“搭乗した環境の運動への適応”においては、ヒューマノイドが搭乗型機器への移乗後に、搭乗型機器のもつ運動の特性に適応するための調速行動制御系とその内部のゲインチューニングを自律的に実行する Mini-batch-based closed-loop Learning Control (MLC) によるゲイン学習法を提案し、搭乗型機器上でヒューマノイドが自身の速度を発散させることなく定位する調速行動を実現した。

本研究で示したこれらの課題解決へのアプローチは、オンライン探索指標の概念により共通に解くことが可能であり、安定的な搭乗型機器ならびに能動的に駆動する機能を備えた搭乗型機器に対する運転操作と適応行動に関して実現可能とするものである。

8.4 今後の課題と展望

本研究の目的は、ヒューマノイドによる搭乗型機器の運転操作ならびに適応行動の独立した行動実現であるが、実際の環境においては運転操作と適応行動が互いに共存し合い、併合

や分離を繰り返している状況である。一例として人間が搭乗型機器を操る場合は、初めて訪れるような路面環境であっても路面の状況により変化する搭乗型機器の動きに適応しながら運転操作を並行して行えている。本研究が提案している行動戦略の処理は、運転操作と適応行動が互いに干渉しない仮定のもとで、既定の課題に対してのみ並行処理や切替処理を実行していることから、運転操作と適応行動が互いに共存する場合や、未知の行動課題が発生した場合などにも対処できるとは言い難い。特に本研究では運転操作の課題を3種類、適応行動の課題を2種類に限定して議論しているが、実際には更なる行動課題も数多く存在しており、尚且つそれらが二重三重に複合した行動課題も存在することから、本研究で提案したアプローチがどこまでの課題を解くことができるかについてまずは判断する必要がある。

また、本研究で対象としている搭乗型機器は非常に限定的であり、事前に人間が搭乗型機器固有のキネマティクス等を見積もった上で、どの軌道を修正するか、オンライン探索指標における探索変数は何に設定すべきか、どの制御系に対する何のゲインを学習対象とするか、など人間が思考・計画するプロセスも多く存在している。

以上のことから、まずは本研究で提案した手法により解決され得る課題に関して、本論文が目的とする「搭乗型機器の運転操作と適応行動」に限らずどこまでの範囲の行動タスクを解くことが可能であるかを吟味し、本研究の手法によりカバーが困難な課題へと汎用性を高めていくことが望まれる。実世界に存在する搭乗型機器への適用についても、図8.1で示したように現状の方法論でカバーされうる領域は限られており、不安定系かつ受動駆動系に属する搭乗型機器へと方法論を一般化させるためには、転倒しないために必要となる即応的なバランス修正に貢献し得る軌道は何か、迅速に適応可能なゲイン学習法とその制御系のゲインは何か、を人間もしくはヒューマノイド自身が思考する過程が重要となる。

搭乗型機器の種類は広がりを持つことから、複数台の搭乗型機器に対する乗降行動や運転操作行動を継続的に組み合わせた、ヒューマノイドによる搭乗型機器の乗り換え行動等は有益な展望課題であるといえる。今後は、そのような展望課題に取り組み、ヒューマノイドの機動性や移動形態の更なる変節へと波及効果を発展させていく。

謝辭

本論文は筆者が東京大学大学院情報理工学系研究科知能機械情報学専攻在学中に、情報システム工学研究室（JSK）において稲葉雅幸教授のご指導のもとで執筆した博士論文です。

稲葉雅幸教授には、本論文の柱には欠かせないヒューマノイドをはじめとするロボティクスの思想や概念について、修士課程から現在に至るまでの多くの時間をかけて、常日頃からご指導頂きました。実世界でロボットは今後何をしていかなければならないか、といったロボティクスの研究に携わることの意義やモチベーションが筆者の研究の原動力となり、ご教授頂いた雄大なロードマップや期待され得る波及効果の考え方が本研究の基盤となっております。特に、本論文のテーマであるヒューマノイドが搭乗型機器を運転し適応することの真に解くべき課題について、筆者が見失いがちな見識を鋭くかつ補助的にご指摘くださりましておかげで本論文を書き上げることができました。ロボットの行動研究を考える指針および行動実現の実験構想に至るまでの広い視野を筆者の成長の糧として享受できましたことに深く感謝申し上げます。

岡田慧教授には、研究の進め方から、ロボットの行動実現を達成するために必要不可欠な知能情報処理、知能システムのソフトウェア開発方針まで、幅広い知見をご指導頂きました。本研究のヒューマノイドによる行動実現では、制御や学習などの知的な情報処理が根幹をなしており、先生からご指導賜りました設計思想の享受によりはじめて実現される方策となっております。また、プログラミングをはじめとするソフトウェアの設計方針にとどまらず、研究を進めていく上での目標や目的を常に意識することの重要性についての的確にご指摘頂きました。その恩恵により、本当に必要とされている機能実現のあり方を見失わずにここまで研究を進めることができましたことを感謝致します。

中村仁彦教授、國吉康夫教授、山本江淮教授には、本論文の審査に関しましてご多忙の折、大変貴重な議論を深めることができましたことを感謝申し上げます。論文の核となるフィロソフィーの重要性について、ご指摘を賜りましたことを深く意識して享受致します。

垣内洋平特任准教授、野沢峻一元特任講師、浅野悠紀助教には、ロボットの扱い方から研究の相談、論文のまとめ方に至るまで常に親身になってご指導頂きました。垣内先生には、DARPA Robotics Challenge (DRC)、ジャパンバーチャルロボティクスチャレンジ (JVRC)、国際ロボット展 (iREX) などの学外での貴重なロボット運用経験を数多く積ませて頂きました。そのご支援のもと、様々な種類のロボットやシステムを扱わせて頂いたことを深くお礼申し上げます。退官された野沢先生には、筆者が修士課程の頃より研究の相談はもちろんのこと、等身大ヒューマノイドロボット“HRP2-JSK”の基本的な動かし方から、その動作原理や

ヒューマノイドシステムの処理方法まで、一から全てのことを昼夜問わずご教授頂きました。ヒューマノイドを扱い始めて右も左も分からない筆者の質問に対しても常に的確に回答して頂き、さらに大きな相談から小さな相談に至るまで親身にアドバイスを賜りました。また、論文のまとめ方や構成などの添削に関しても丁寧にご指導くださりましたことに深く感謝の意を表します。浅野先生には、筆者が研究室に配属されてから最初に所属した腱駆動班で、腱駆動ヒューマノイドロボット“腱志郎”の動かし方をはじめ、他のロボットにはない腱駆動ロボットの特性や能力、魅力について手厚くご指導頂きました。腱駆動ロボットの修理から、修士課程での研究の進め方の相談、筆者が初めて執筆する投稿論文の書き方のアドバイスや添削など、様々な局面で広くお世話になりました。

本論文のヒューマノイドによる行動実現に貢献して下さった、ご卒業された上月豊隆博士、白井拓磨特任助教、卒業生の溝花弘登君にも深く感謝申し上げます。研究の議論から食事に至るまで共に研究生活を過ごしてきた、HRP-2班の方々、腱駆動班の方々、そして工学部2号館73B1号室、73B2号室、工学部8号館326・330号室、602・610号室でお世話になったの方々、および同期の方々、にも厚く感謝を申し上げます。これまでJSKを創ってきた先生方、卒業生・現役生の方々にも感謝致します。

そして、本論文の実証実験において、壊れることがありながらも最後まで行動実現の役目を全うした、等身大ヒューマノイドロボット“HRP2-JSK”（HRP-2 7号機）、腱駆動ヒューマノイドロボット“腱志郎”，および実験の対象となった搭乗型機器とその設計・開発者様方にお礼申し上げます。特にHRP-2の修理対応・部品提供・メンテナンス・保守点検をはじめ、手厚く幅広いサポートをして下さったカワダロボティクス株式会社 影山様、後藤田様に感謝の意を表します。

最後に、長き研究生活の道のりを歩むにあたって、経済面、精神面などで常に支え続けてくれた家族に心より感謝致します。

多くの方々の支えにより、本論文が完成しましたことに改めて感謝の意を表します。

2018年6月8日 木村 航平

付録

A.1 搭乗型機器の構成

A.1.1 三輪車の仕様

搭乗型機器の対象となる三輪車の寸法（図 A.1，図 A.2）および構成について述べる。三輪車の三輪構造は1ステアリングホイールかつバックホイールが2輪となっている。2輪のバックホイールのうち、左側のホイールが駆動輪であり、対向の右側が受動車輪となっている。また、三輪車のハンドルの最大操舵角は $90[\text{deg}]$ 程度である。ヒューマノイドが搭乗して運転する場合においては、3.4.2 項で述べた通り、ヒューマノイドの自己干渉制約から最大操舵角は $90[\text{deg}]$ 未満となる。ペダルについてはビンディング型式のものでなく、日常的なシティサイクル型を本研究では採用する。クランク1回転に対する走行距離 D_c が $2.45[\text{m}]$ となるようなクランクギアに設定している。等身大ヒューマノイドロボット HRP2-JSK [110] が接触する箇所となるハンドルグリップ、ペダル、サドルは図 A.3 であり、それぞれゴム製、プラスチック製、ウレタン製のものを使用している。

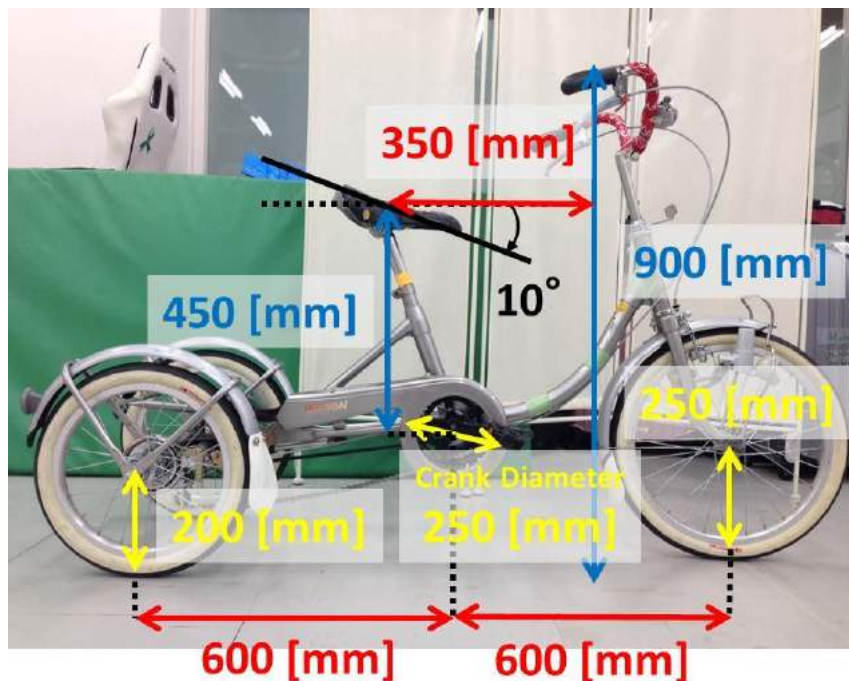


図 A.1: Side View of Tricycle

Crank Shaft to Front Shaft: $600[\text{mm}]$, Crank Shaft to Back Shaft: $600[\text{mm}]$, Crank Diameter: $250[\text{mm}]$, Front Tire Radius: $250[\text{mm}]$, Back Tire Radius: $200[\text{mm}]$, Saddle Height from Crank Shaft: $450[\text{mm}]$, Handle Grip Height from Ground: $900[\text{mm}]$, and Horizontal Distance from Saddle to Handle Grip: $350[\text{mm}]$

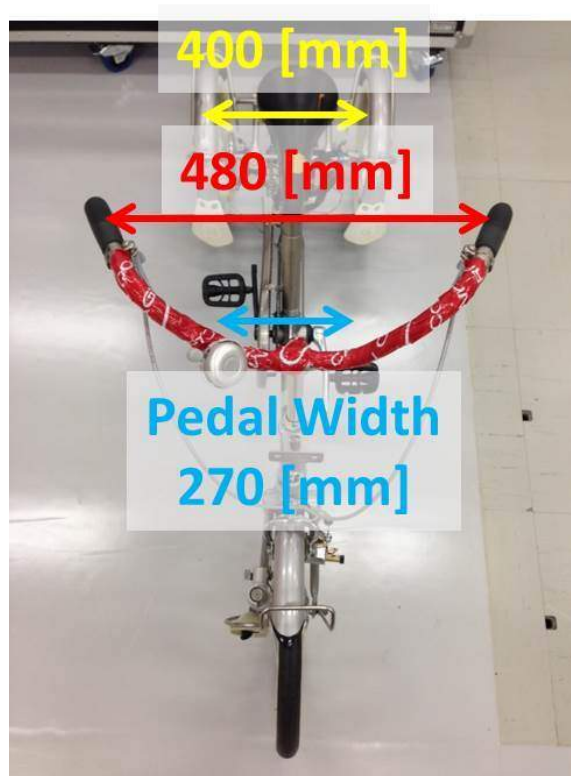


図 A.2: Top View of Tricycle

Pedal Width: 270[mm], Handle Width: 480[mm], and Back Wheel Tread Width: 400[mm]

A.1.2 平行二輪電動スクーターのモデル化

搭乗型機器の対象となる平行二輪電動スクーターの概要について述べ、最終的に平行二輪電動スクーターの操作量と搭乗者となるヒューマノイドに出力される速度および Yaw まわりの旋回角速度の関係のモデル化を行う。

平行二輪電動スクーターの駆動原理とボード傾斜角-車輪角速度の関係

本研究で使用する平行二輪電動スクーターは、Vivo Market 株式会社より販売されている“V-BOARD” [58] と呼ばれる製品であり、電源投入後、図 A.4 に示すように 2 面のボードに両足を乗せた状態で左右の足首を Pitch 方向に動かして調整することで、その場での停止や前進・後退・方向転換が可能である。この操作方法はパーソナルモビリティとして一般的に



図 A.3: Left: Handle Grip Rubber of Tricycle, Right: Plastic Pedal of Tricycle, and Bottom: Urethane Saddle of Tricycle

普及しているハンドル付きのセグウェイ [192] の駆動原理と少し異なり，セグウェイ [192] は両足を乗せるボードが 1 面のみで構成されており，搭乗者の重心移動により 2 輪を駆動させて運転する方式であることが知られている [193][194][195]．また，方向転換に関してもハンドルを両手で傾げることで旋回を行う方式となる．一方で，平行二輪電動スクーターの場合はフォースプレートのような搭乗者の重心位置や足裏の圧力中心を計測してフィードバックする装置は組み込まれておらず，左右のボード上の接触センサによる二値の接触/非接触の判定のみにより，左右の車輪の制御駆動モードを ON/OFF で切り替える構造となっている．したがって，搭乗者の前進・後退・旋回・停止などの運動はすべて左右のボードを傾斜させることによる独立した 2 輪の駆動操作により行う．

左右の足首の Pitch 方向の操作量，すなわち平行二輪電動スクーターの左右に位置する 2 面のボードの Pitch 傾斜角と，2 つの車輪の回転角速度の駆動関係について観測する．平行二輪電動スクーターの片側のボードの傾斜角，および片側の車輪の回転角速度を計測する実

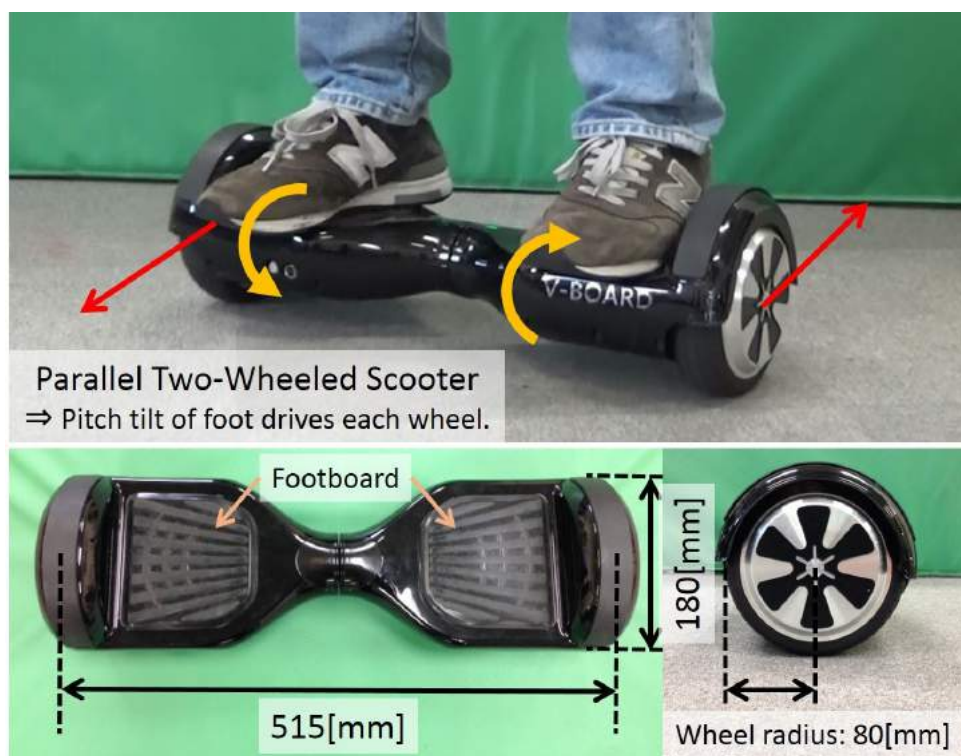


図 A.4: Parallel Two-Wheeled Electric Scooter “V-BOARD”

験構成を図 A.5 に示す。左右のボードのうち一方（図 A.5 では進行方向左側）を固定させて、反対の計測側のボード（図 A.5 では進行方向右側）を傾斜させて観測する実験作業を左右のボード各々で行う。SmartTool Technologies 社製の Digital Protractor Pro 3600 を使用し、ボードの傾斜角度 [deg] のデジタル値を読み取る。このときの車輪の回転速度 [r/min] を小野測器株式会社製のデジタルタコメーター HT-5500 から取得し、単位換算により角速度 [rad/s] を求める。デジタルタコメーターは非接触式であり、回転体である車輪に直接接触することなく赤色光を照射させて車輪側面に貼付した反射マーカからの反射光を検出することで回転速度を計測する。反射光の検出可能距離は 20[mm]~300[mm] であり、本実験構成ではデジタルタコメーターの光源から反射マーカまでの距離を 80[mm] としている。また、車輪を駆動させるためには電源投入に加えて制御駆動モードを ON にする必要がある。ボード上の接触センサの判定が常に計測中は「接触」となるように接触部をベルトで締めて負荷を与えている。

以上の実験環境を構築した上で、平行二輪電動スクーターのボードの傾斜角（右側を

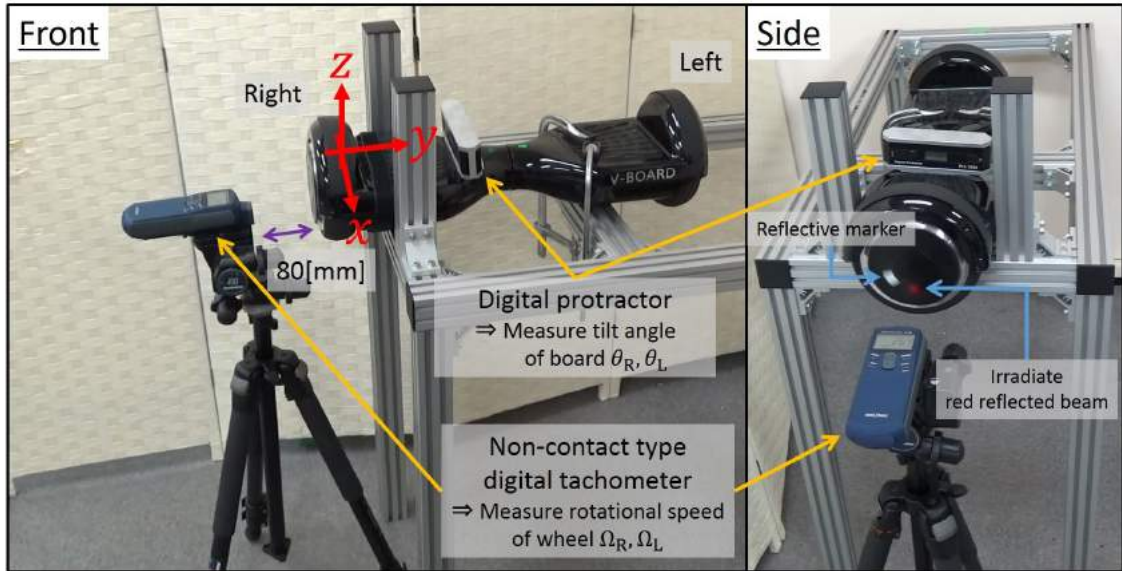


図 A.5: Experiment to Measure the Rotational Angular Velocity of “V-BOARD” Wheel
The tilt angle [deg] of each board is measured by the digital protractor, and the rotational speed [r/min] of each wheel is measured by the non-contact type digital tachometer.

θ_R [deg], 左側を θ_L [deg] とする) を 0.5[deg] 刻みで設けたときの車輪が出力する回転角速度 (右側: Ω_R [rad/s], 左側: Ω_L [rad/s]) の計測値をプロットした結果が図 A.6 の赤点 (θ_R, Ω_R), 青点 (θ_L, Ω_L) のデータである. ここで, $\theta_R, \theta_L, \Omega_R, \Omega_L$ はすべて図 A.5 で示した y 軸まわりを正とする. (θ_R, Ω_R) と (θ_L, Ω_L) の組は共に θ_R, θ_L が約-17[deg]~13[deg] の範囲では比例関係が観察され, その範囲を超えるとボードの傾斜角に依らず車輪の回転角速度は一定に落ち着いている. 目的とする调速行動 (行動レベル3) においては停止が目標であり, 停止動作の特性としては $\Omega_R = 0, \Omega_L = 0$ 近傍での議論が主となるため, 前者の比例関係の特性に注目する. 本論文では调速行動の目標速度を 0 (停止) としているが, 任意に設定した目標速度での调速行動に取り組む場合は, 後者の比例関係の特性範囲外の飽和領域についても考慮する必要があることに留意する. 飽和領域となる前後 8 点を除いた計測点 (θ_R, Ω_R), (θ_L, Ω_L) のデータ群に対して, 最小二乗法を利用して線形回帰した結果が図 A.6 の黄線, 緑線となる. このときの線形回帰の相関係数の二乗値 R^2 を計算すると, 黄線, 緑線それぞれ $R^2 = 0.9993, R^2 = 0.9990$ であり, ほぼ 1 に近い値となっていることから一次関数としてフィッティングするのに妥当であることが観察される. 以上から, 比例係数 K_R, K_L と θ_R, θ_L 切片のオフセットとなるボード傾斜角 θ_{RO}, θ_{LO} を用いて, $[\Omega_R \ \Omega_L]^T$ は次式のようにモデル化できる.

$$\begin{bmatrix} \Omega_R \\ \Omega_L \end{bmatrix} = \begin{bmatrix} 4.6203 & 0 \\ 0 & 4.3356 \end{bmatrix} \begin{bmatrix} \theta_R + 1.7871 \\ \theta_L + 1.9972 \end{bmatrix} = \begin{bmatrix} K_R & 0 \\ 0 & K_L \end{bmatrix} \begin{bmatrix} \theta_R - \theta_{RO} \\ \theta_L - \theta_{LO} \end{bmatrix} \quad (\text{A.1})$$

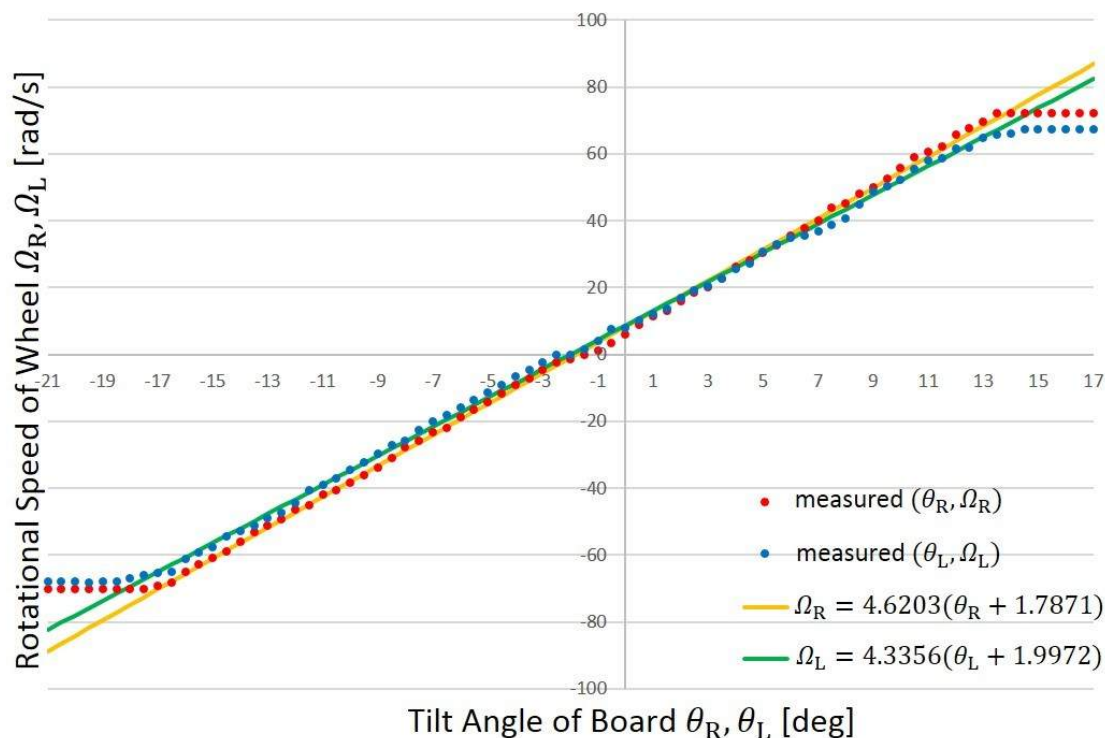


図 A.6: Relation between the Tilt Angle of Board θ_R, θ_L [deg] and the Rotational Speed of Wheel Ω_R, Ω_L [rad/s]

The proportionality (yellow line, green line) between θ_R, θ_L and Ω_R, Ω_L is obtained from the linear regression for the measured data (θ_R, Ω_R) (red dots) and (θ_L, Ω_L) (blue dots).

平行二輪電動スクーターの速度と旋回角速度

前節の式 A.1 より，平行二輪電動スクーターのボード傾斜角の入力から車輪の回転角速度の出力への駆動関係のモデル化を行った．本節ではこれに基づき，平行二輪電動スクーター車体の二次元平面上における速度と Yaw まわりの旋回角速度の出力への移動要素のモデル化を行う．

式 A.1 から両輪の回転角速度が得られたことで、車輪が床面を滑ることなく回転するとき、平行二輪電動スクーターの両端点での速度 $[V_R \ V_L]^T$ [m/s] は以下となる。

$$\begin{bmatrix} V_R \\ V_L \end{bmatrix} = r \begin{bmatrix} \Omega_R \\ \Omega_L \end{bmatrix} \quad (\text{A.2})$$

ここで r [m] は車輪の半径である。

また、図 A.7 は平行二輪電動スクーターの車体相対でみたときの二次元平面上（床面上）における速度 v [m/s] と旋回角速度 ω [rad/s] の図示であり、車体相対の基準座標を平行二輪電動スクーターの中心とする。 W [m] は平行二輪電動スクーターの車体幅である。

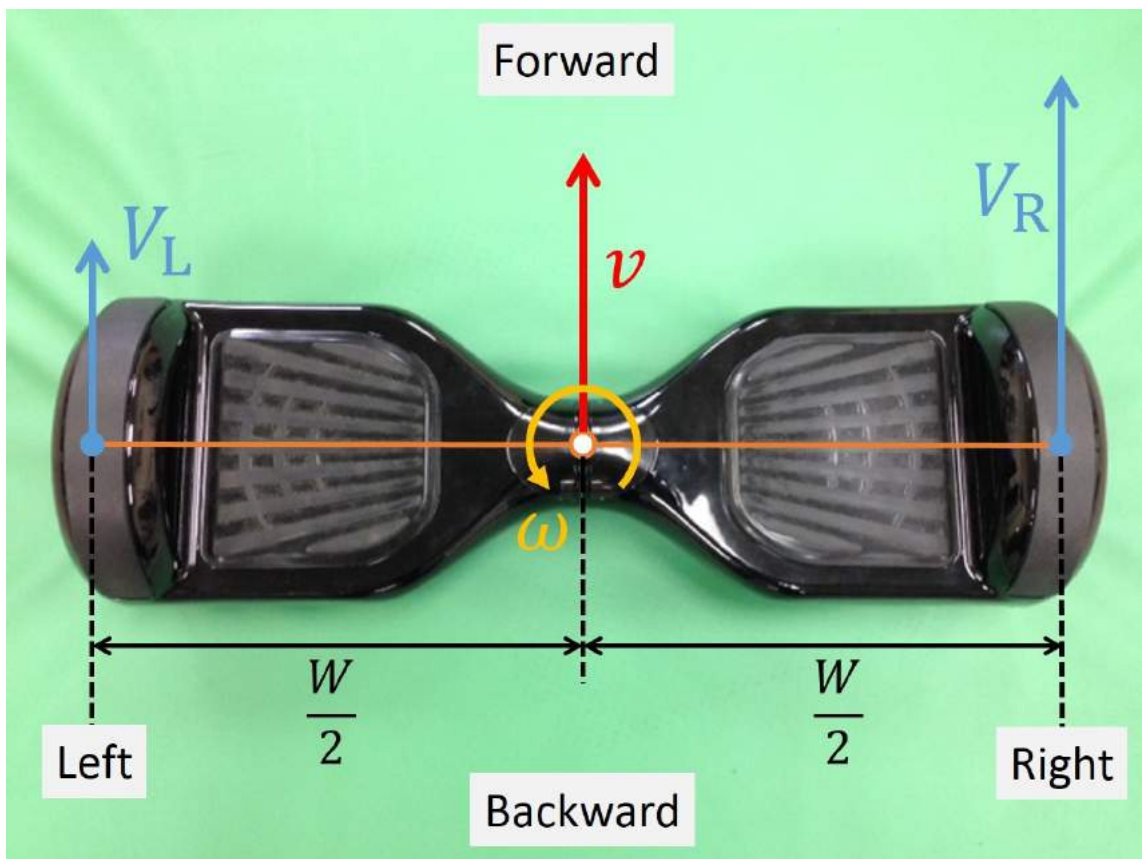


図 A.7: Geometric Relationship between the Linear Velocity $[V_R \ V_L]^T$ at the Endpoints of Parallel Two-Wheeled Electric Scooter and the Linear/Angular Velocity $[v \ \omega]^T$ of Vehicle Body

図 A.7 より平行二輪電動スクーターの端点速度 $[V_R \ V_L]^T$ と車体速度・旋回角速度 $[v \ \omega]^T$

の関係は次式で表すことができる.

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{W} & -\frac{1}{W} \end{bmatrix} \begin{bmatrix} V_R \\ V_L \end{bmatrix} \quad (\text{A.3})$$

式 A.1, 式 A.2, 式 A.3 より $[v \ \omega]^T$ は以下で与えられる.

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = r \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{W} & -\frac{1}{W} \end{bmatrix} \begin{bmatrix} K_R & 0 \\ 0 & K_L \end{bmatrix} \begin{bmatrix} \theta_R - \theta_{R0} \\ \theta_L - \theta_{L0} \end{bmatrix} \quad (\text{A.4})$$

式 A.4 は平行二輪電動スクーターの左右の操作傾斜角 θ_R, θ_L を入力としたときの搭乗者 (ヒューマノイド) の速度 v と旋回角速度 ω の出力との関係を表している. このように車体の傾きに応じて, 車輪の角速度を指令することでバランスをとる倒立振り型の制御方式は文献 [196] でも提案されている.

A.1.3 電動一輪車のモデル化

搭乗型機器の対象となる電動一輪車の概要について述べ, 搭乗者であるヒューマノイドの目標 ZMP と出力される電動一輪車の速度の関係をモデル化する. 本研究で用いる電動一輪車は, 株式会社オオトモが販売している “Ninebot One S2” [191] という機体であり, 図 A.8 に示すように 1 輪の電動駆動する車輪と搭乗者の足部を乗せるためのボードを備えている. A.1.2 項で述べた平行二輪電動スクーターの 2 つのボードと比較して, ボード単体には傾斜検知の用途のセンサ類は搭載されておらず, 車体内部に 1 輪の車輪を制御して Pitch 方向の傾斜を立て直す働きをもつ制御系が導入されている. したがって, 搭乗者の Pitch 方向の傾きは車輪を回転させて車体速度として出力することで水平に保持する機能を有するが, Roll 方向の傾きに関しては制御する機能を備えてなく, 搭乗者自身がバランスを制御する他, 図 A.8 -(II) のように左右の足部ボードの下面に転倒防止用の補助輪 (受動輪) をハードウェアとして取り付ける方法がある. 補助輪を取り付けることで, 図 A.8 -(I) のような Roll 方向の傾きを介助する必要なく車体を立てることが可能である.

補助輪を有した場合の “Ninebot One S2” (図 A.8 -(II)) は車体自体は安定であるが, 搭乗者 (ヒューマノイド) の重心, ZMP に応じて 1 輪の車輪が電動駆動されるため, それにより速度も出力される状態である. 図 7.8 で示したシステム内の最終目標 ZMP p^{d**} が “Ninebot One S2” にかかる時の車体に出力される速度 (表 7.3 で定義した v_{traj}) の関係を表したモデル図を図 A.9 に示す. 図 A.9 に従い, 最終目標 ZMP p^{d**} と出力される並進速度 v_{traj} の関係を導く.



(I) “Ninebot One S2” without Training Wheels

(II) “Ninebot One S2” with Training Wheels

図 A.8: Electric Unicycle “Ninebot One S2”

はじめに，“Ninebot One S2”の中心に作用する Pitch 方向（y 軸まわり）のモーメントに応じて車輪が回転すると仮定して，車体におけるモーメントの観測点（ゼロ点） p^0 での発生モーメント M_y^0 を求める． p^{d**} はゼロモーメントポイント，すなわち搭乗者であるヒューマノイドの圧力中心として捉えることができ，“Ninebot One S2”に対してヒューマノイドが静的姿勢である場合， p^{d**} には鉛直下向きにヒューマノイドの自重 W が掛けられることに等しく，このときの p^0 まわりの発生モーメント M_y^0 は以下で記述される．

$$M_y^0 = W(p^{d**} - p^0) \quad (\text{A.5})$$

次に，発生するモーメント M_y^0 に比例して車輪の回転速度 Ω が出力されるとすると，

$$\Omega = kM_y^0 \quad (\text{A.6})$$

の関係であり，さらに車輪が滑ることなく地面を転がり走行する場合，

$$v_{traj} = r\Omega \quad (\text{A.7})$$

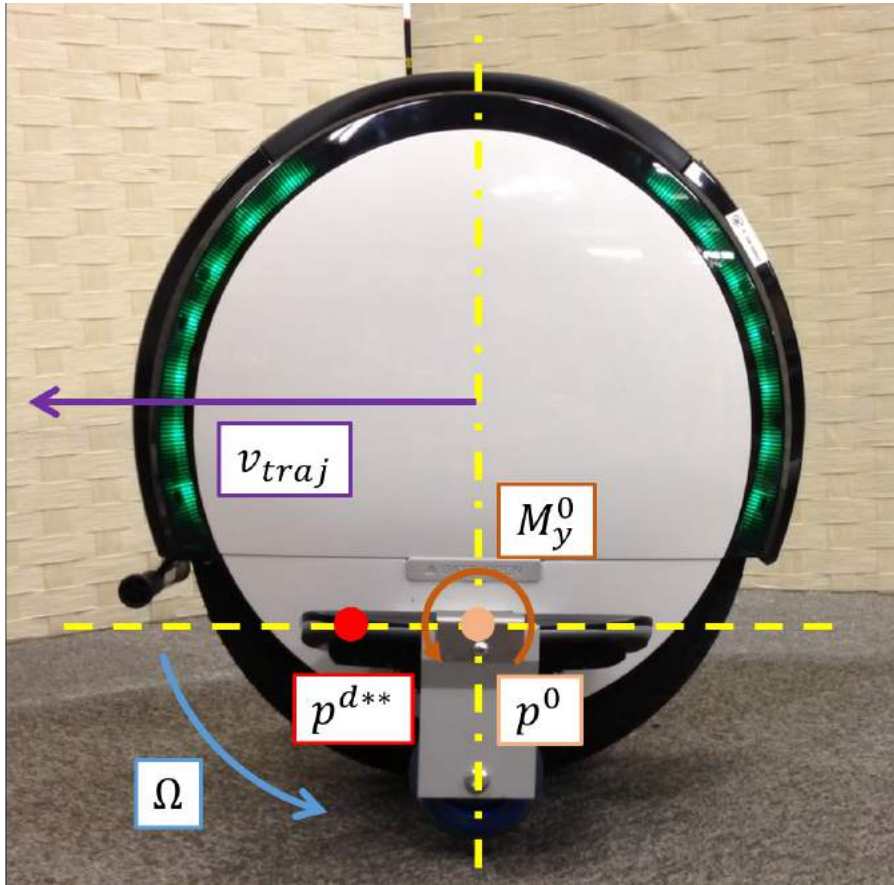


図 A.9: Parameters about Linear Velocity Modeling of “Ninebot One S2”
Reference ZMP of humanoid robot is put on the footboard of “Ninebot One S2”.

の関係がある。ここで、 k は比例定数、 r は車輪の半径を表す。式 A.5，式 A.6，式 A.7 より p^{d**} と v_{traj} の関係式は以下となる。

$$v_{traj} = rkW(p^{d**} - p^0) \quad (\text{A.8})$$

$K_0 = rkW$ として定数項をまとめて p^0 も定数となることに留意すれば、式 A.8 の両辺を時間微分しラプラス変換を行うと、周波数領域で p^{d**} と v_{traj} の関係は以下で表される。

$$v_{traj} = K_0 p^{d**} \quad (\text{A.9})$$

上記のモデルは図 7.8 の併合システム内においても導入されている。

A.2 ヒューマノイドロボットの構成

A.2.1 等身大ヒューマノイドロボット“HRP2-JSK”

ハードウェア構成

本研究で示す行動実現の実機検証で用いる等身大ヒューマノイドロボットは、ヒューマノイドロボット HRP-2 [197][198][199] を独自に改良した“HRP2-JSK” [110] である (図 A.10)。

HRP-2 の特徴としては、腰軸に 2 自由度を持つなど他の等身大ヒューマノイドロボットと比較し多自由度であること、メンテナンス製、安全性を考慮したその完成度の高さ等があげられる。詳細なハードウェア仕様は川田技報 [13] を参照されたいが、本節では当研究室で改良を加えている点について説明する。

HRP2-JSK のハードウェアデータを表 A.1, 表 A.2 に示す。表 A.1 の寸法データ (Height, Width, Depth の各項) はロボットの構造を記述した CAD データを元に算出した値である。Weight データは、電子計りで実測した値を用いている。DoF データは関節の駆動軸数である。表 A.2 は川田技報のデータを元に追加および修正した各関節仕様である。

表 A.1: Hardware Data of HRP2-JSK [200]

Height [mm]	1607
Width [mm]	627
Depth [mm]	324
Weight [kg]	56.5
DoF	32

ハードウェア構成に関する具体的な変更点は次の通りである。

1. より複雑な動作を可能にするために手首の自由度を増やし、腕の自由度を 6 自由度から 7 自由度に変更。
2. 襟と頬の干渉を避け首の可動範囲を広げるため、首を従来より 30[mm] 伸長。
3. 視覚処理を用いて障害物認識, 周囲環境認識, 自己位置認識を行うために、標準搭載のカメラデバイスとは異なるステレオカメラおよび LRF (Laser Range Finder) を搭載。頭部には、障害物認識, 周囲環境認識のための点群データを取得する用途のステレオカメラ“MultiSense S7”と Hokuyo LRF (Laser Range Finder) を統合した“MultiSense



☒ A.10: Life-Sized Humanoid Robot “HRP2-JSK”

表 A.2: Joint Specification of HRP2-JSK (Modified from [13])

部位区分 総自由度数 32	関節名称		可動範囲 [deg]	最終出力	
				最大トルク [Nm]	最大回転数 [rpm]
頭部 自由度数 2	首関節	ピッチ	前方 45, 後方 30	9.6	64.0
	首関節	ヨー	左右方向 45	9.6	64.0
腕部 自由度数 $8 \times 2 = 16$	肩関節	ピッチ	前方 180, 後方 60	160.3	39.2
	肩関節	ロール	外側 95, 内側 10	100.4	22.9
	肩関節	ヨー	外側 92, 内側 92	48.7	39.4
	肘関節	ピッチ	前方 137, 後方 2	108.5	21.2
	手首関節	ヨー	外側 92, 内側 92	13.9	44.0
	手首関節	ロール	前方 90, 後方 90	18.7	63.4
	手首関節	ピッチ	前方 90, 後方 90	16.0	26.5
	ハンド親指	ロール	開方向 60, 閉方向 16	13.9	44.0
腰部 自由度数 2	腰関節	ピッチ	前方 60, 後方 5	278.9	22.8
	腰関節	ヨー	左右方向 45	151.8	41.9
脚部 自由度数 $6 \times 2 = 12$	股関節	ピッチ	前方 125, 後方 42	151.4	42.1
	股関節	ロール	外側 35, 内側 20	175.4	36.2
	股関節	ヨー	外側 45, 内側 30	86.2	33.8
	膝関節	ピッチ	前方 2, 後方 150	168.3	37.9
	足首関節	ピッチ	前方 75, 後方 42	131.6	48.3
	足首関節	ロール	前方 20, 後方 35	73.1	87.0

SL” [155] に加えて、ヒューマノイドの自己位置認識のためのビジュアルオドメトリデータを取得する用途のステレオカメラ “ZED” [183] をそれぞれ構成 (図 A.11)。

4. 計算能力向上のため、体内の 2 つの組み込み型 PC の CPU ボードを変更。制御用リアルタイム PC および画像処理用 PC は、PentiumIII 対応から Core2Duo 対応へと変更。
5. 体内搭載の組み込み型 PC に加えて、ステレオカメラ “ZED” によるビジュアルオドメトリ計測の向上を図る GPU 付属の PC “ZBOX”¹⁾ を背中に外付け。“ZED” は USB3.0 により “ZBOX” と接続。
6. 手首の可動範囲の拡大。
7. CPU ファンを静音かつ出力の高いものへと変更。また冷却能力向上のために、モータドライバまわりに冷却ファンを追加。

¹⁾https://www.zotac.com/product/mini_pcs/all

8. 通気性向上のため、背中側カバーに吸排気口を追加.
9. 拡張性向上のため USB, 5V, 12V, 48V ラインを体内 PC から両手先, 足先, 臀部へと延長.
10. HRP2-JSK は 1 軸ハンドに把持安定化を狙った爪 (図 A.12 -(a)) を搭載.
11. アルミニウムフレームおよび CFRP (Carbon Fiber Reinforced Plastics), 緩衝用ウレタンから構成される臀部カバー (図 A.12 -(b)) を搭載.

特に、三輪車のハンドルグリップ等を把持するための HRP2-JSK のハンドは、1 自由度グリップ型ハンドであり、把持力に優れている。また、相対する指リンクには把持安定性向上のため、アルミ板を加工した爪 (図 A.12 -(a)) を取り付けしている。臀部には、搭乗時のサドルからの反力を緩衝するためのカバーが備えられており、アルミニウムの支柱の先に CFRP カバーとウレタンが外装として機能している。

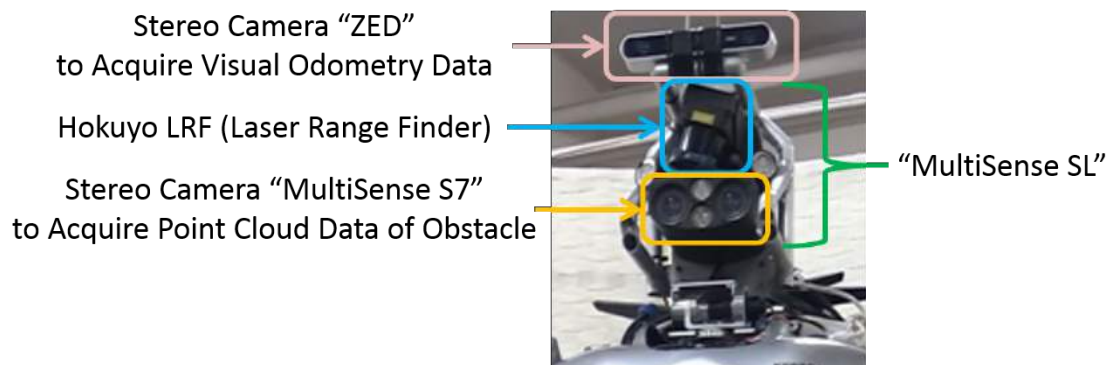


図 A.11: Visual Sensors Mounted on HRP2-JSK Head



(a) Aluminium Craw Attached to Hand

(b) Hip Cover around Crotch

図 A.12: Hardware Improvements of HRP2-JSK

ソフトウェア構成

HRP2-JSKのソフトウェア構成の概略図を三輪車の運転操作の場合を例に挙げて図 A.13 に示す。

図 A.13 における各コンポーネントについて以下で概説する。

ROS

ROS はオープンソースのロボット関連ソフトウェア群であり、通信ライブラリ、ツール、基盤、エコシステムから構成される。統合システムであるロボットの複雑高度な要素機能を、様々な環境で共有することでロボット開発の効率を向上させることを目標としており、世界中の利用者からなるコミュニティによって支えられている [141]。

ROS は `roscore` と呼ばれるネームサーバを中心として `node` と呼ばれるプロセス間が `message`, `service`, `topic` といった形式で相互に通信し合うことによって、分散処理を実現している。本システムでは、動作計画、視覚認識、ユーザインタフェースといった上層の処理の一部に、ROS の公開パッケージを利用することで新規部分に注力し開発効率を高めている。主なプログラムとしては、オブジェクト指向、幾何モデラなどの機能を有する Lisp 拡張言語 `euslisp`²⁾ [201] に、ROS の機能をもたせた `roseus`³⁾ をインタラクティブなプログラミングの

²⁾<http://wiki.ros.org/euslisp>

³⁾<http://wiki.ros.org/roseus>

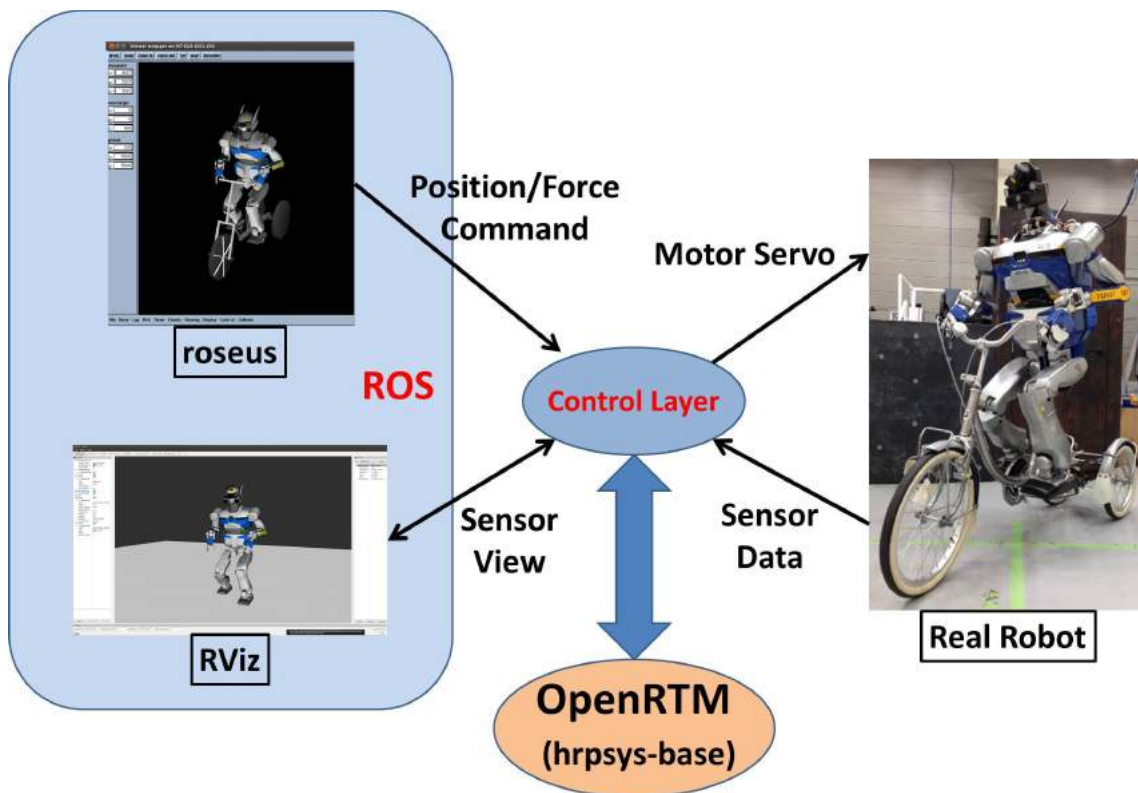


図 A.13: Software System Configuration for HRP2-JSK with Tricycle

インタフェースとして使用している。また、RViz⁴⁾と呼ばれるロボット状態可視化ツール、rosviz⁵⁾と呼ばれるロボットログ管理ツール等により、ロボット実験ソフトウェアの根幹部分を構成している。

OpenRTM

OpenRTM [164] は産業技術総合研究所が開発・配布するロボットシステムのためのソフトウェアプラットフォームである。OpenRTMではデータポートと呼ばれる形式によりデータ送受を行うことで、RT-Componentと呼ばれる単機能のプログラムを接続しロボットシステムを構成する。

⁴⁾<http://wiki.ros.org/rviz>

⁵⁾<http://wiki.ros.org/rosviz>

hrpsys-base はハードウェアに依存するソフトウェアをひとつの RT-Component (RobotHardware) に分離し, 残りのコンポーネントでは関節角補間 (SequencePlayer), バランス制御 (AutoBalancer), 安定化制御 (Stabilizer), インピーダンス制御 (ImpedanceController) といったロボット共通の制御を行う. 括弧内は hrpsys-base における各 RT-Component の名称である. ロボットのハードウェア構造記述には VRML と呼ばれるフォーマットのロボットモデルを使用する.

図 6.6 で示した调速行動制御系を例に挙げ, ROS と OpenRTM が共存したリアルタイム制御系の通信構成図を図 A.14 に示す. ロボットに搭載されたステレオカメラ “ZED” や IMU などのセンサ類, Joystick コントローラなどの外部操縦インタフェースから送信されるデータを ROS の topic として通信し, ZBOX や体内の画像処理用 PC (Vision PC) で処理されたデータ (ロボットの並進速度 v や速度・角速度指令 v^d, ω^d) をデータポート形式で制御用リアルタイム PC (Control PC) に送信する. Control PC では hrpsys-base に基づく Stabilizer コンポーネントのリアルタイム処理により, ロボット実機に対する制御周期 4[ms] 程度の実時間制御を実行している.

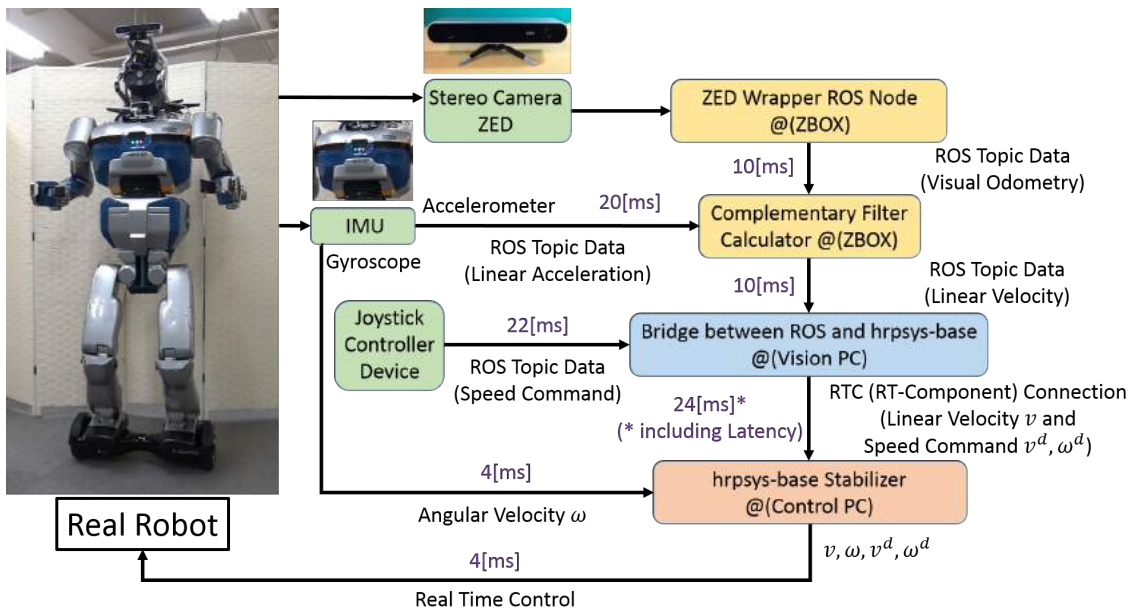


図 A.14: Software System Configuration for Speed Governing Behavior Control System
Purple-colored values represent sampling period.

EusLisp

EusLisp は幾何モデリング機能をもったマルチスレッドなオブジェクト指向の Lisp 言語であり [202], OpenGL を利用した 3 次元表示機能なども持っている. C/C++ 言語に比べ, キーワード引数, プログラムやデータの S 式による表記, lambda 式, ハッシュ表など優れた機能を持つ. また, EusLisp は他言語インタフェースと呼ばれており, C 言語のライブラリを組み込むことができる機能を持っている. これにより, 画像処理, シミュレーションなどを同一の環境で行なうことができる.

実世界に対応した幾何モデル環境を作成し, 実世界とモデル環境との間の整合性を保つことによってモデルベースなアプローチを実現している. EusLisp によりロボットモデルを用いて任意の動作を生成し, 最終的にモデル上のロボットと実ロボットの関節角度列レベルで交互にやり取りすることにより, 実世界とモデル世界の連携を行う. また, 指令先をシミュレータに切り替えることも容易であり実機を用いずとも開発可能である. 本研究においても, オフライン探索指標をベースとした能動機能や重心軌道の計画・生成等の運動学, マスプロパティ情報処理を EusLisp により記述し行動実現を行っている.

A.2.2 腱駆動ヒューマノイドロボット“腱志郎”

ハードウェア構成

本研究で用いる腱駆動ロボットは, 東京大学大学院情報システム工学研究室が開発した筋骨格腱駆動ヒューマノイドロボット“腱志郎”[138][139][203][204][205] と名付けられた人体模倣設計に基づく柔軟かつ多自由度な特徴を有するヒューマノイドである (図 A.15).

腱志郎のハードウェアデータを表 A.3 に示す. 表 A.3 のデータは文献 [138][139][203][204][205] に基づいて整理された腱志郎全身の寸法データ (Height, Width, Depth), 重量データ (Weight), 関節自由度数データ (DoF of Joints), 駆動筋数データ (Muscle Actuators) をそれぞれ表す.

表 A.3: Hardware Data of Kenshiro

Height [mm]	1600	Weight [kg]	51.9
Width [mm]	372	DoF of Joints	64
Depth [mm]	211	Muscle Actuators	87



図 A.15: Musculoskeletal Tendon-Driven Humanoid Robot “Kenshiro” [138][139][203][205]

詳細な構成仕様のデータは文献 [138][139][203][204][205] に譲るとして、本節では本論文で示した行動実現に関して留意する特徴を以下に述べる。

1. 全身の各リンク（骨格に相当）上に配置された maxon の 100[W] ブラシレス DC (BLDC) モータによる、リンクに取り付けられたダイニーマ（筋に相当）の巻き取りにより関節の腱駆動を実現。
2. モータドライバ基板として、高出力アクチュエータ駆動用小型モータドライバ FPGA-01D [203] を採用。
3. 筋となるダイニーマには、人体の腱を模倣した機械的な非線形バネ要素 NST (Nonlinear Spring Tensioner) [206][207] を搭載。
4. 筋の張力を計測する用途で、モータユニット内部に超小型圧縮ロードセル（東洋測器株式会社製，MCDW-50L， $\phi 7$ ，定格 500[N]）を配置 [203]。
5. 足部には CAN 通信規格の 6 軸力センサ（Istituto Italiano di Tecnologia 製）を搭載。
6. 頭部内部に小型ブレイン PC “Intel® NUC”⁶⁾ を搭載することで体内完結型の身体構成を実現。
7. ヤコビアンを介した筋長制御による関節角度，エンドエフェクタ位置の制御が可能。また，筋の張力を計測することで張力制御のモードが利用可能。

ソフトウェア構成

腱志郎のソフトウェア構成を文献 [203] に基づき図 A.16 に示す。図 A.16 のソフトウェアシステムは、ブレイン層，中間層，ロボットボディ層から構成されるレイヤーの階層的構造を形成している。

本研究で示す腱志郎による三輪車の運転操作行動実現（3.5.5 項）では，能動機能の脚と受動機能の脚を図 A.17 に示す協調システムに従って制御する。本節では，レイヤーの各機能の詳細 [203] をもとに，図 A.17 のシステムを対象とした場合のそれぞれの層の働きについて次に整理する。

ブレイン層

ロボットの運動学モデルに基づいた能動機能と受動機能それぞれに対する脚のエンドエフェクタ位置 p_i, q_i の動作計画を行う層。

⁶⁾<https://www.intel.com/content/www/us/en/products/boards-kits/nuc.html>

中間層

下層に位置するロボットボディ層から受け取るセンサデータ（図 A.17 の協調システムでは，受動機能側の脚に搭載された 6 軸力センサから受信する外力変化 Δf_i ）を取得し，上層に位置するブレイン層にそのデータを受け渡す働きをもつ層．各アクチュエータの制御においては，ブレイン層で計画した目標指令値（ $i = i$ から $i = i + 1$ までのシーケンス）を補間した上で，下層のロボットボディ層の制御基板，モータドライバ等の処理系へと伝達・通信する働きを有する．

ロボットボディ層

ロボットの身体に分散配置された制御基板やモータドライバを含むモータユニット，およびセンサデバイスをはじめとするファームウェアレベルの処理を実行する層．センサデータは A/D 変換され中間層へと伝達される．

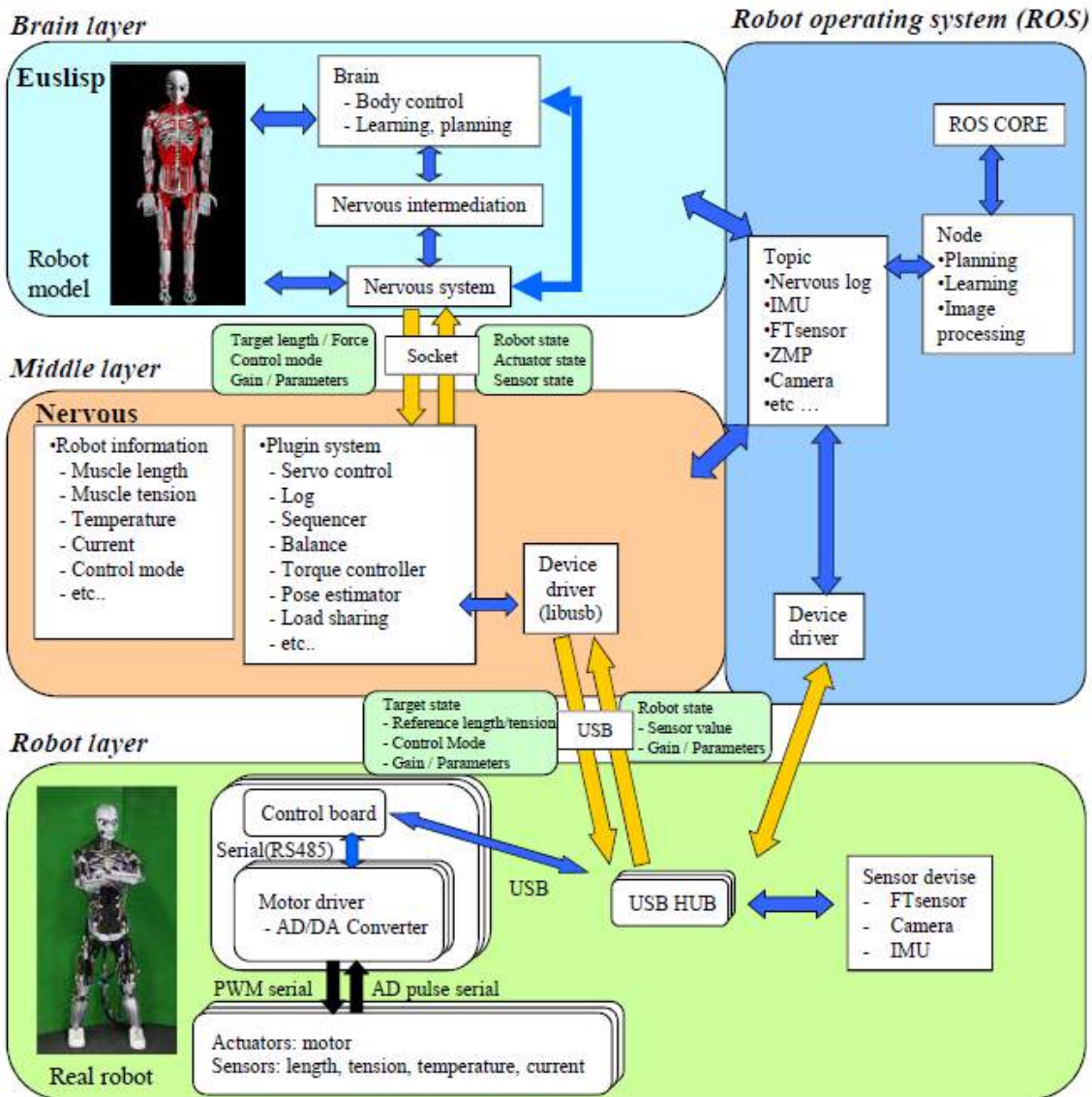


図 A.16: Software System Configuration for Kenshiro [203]

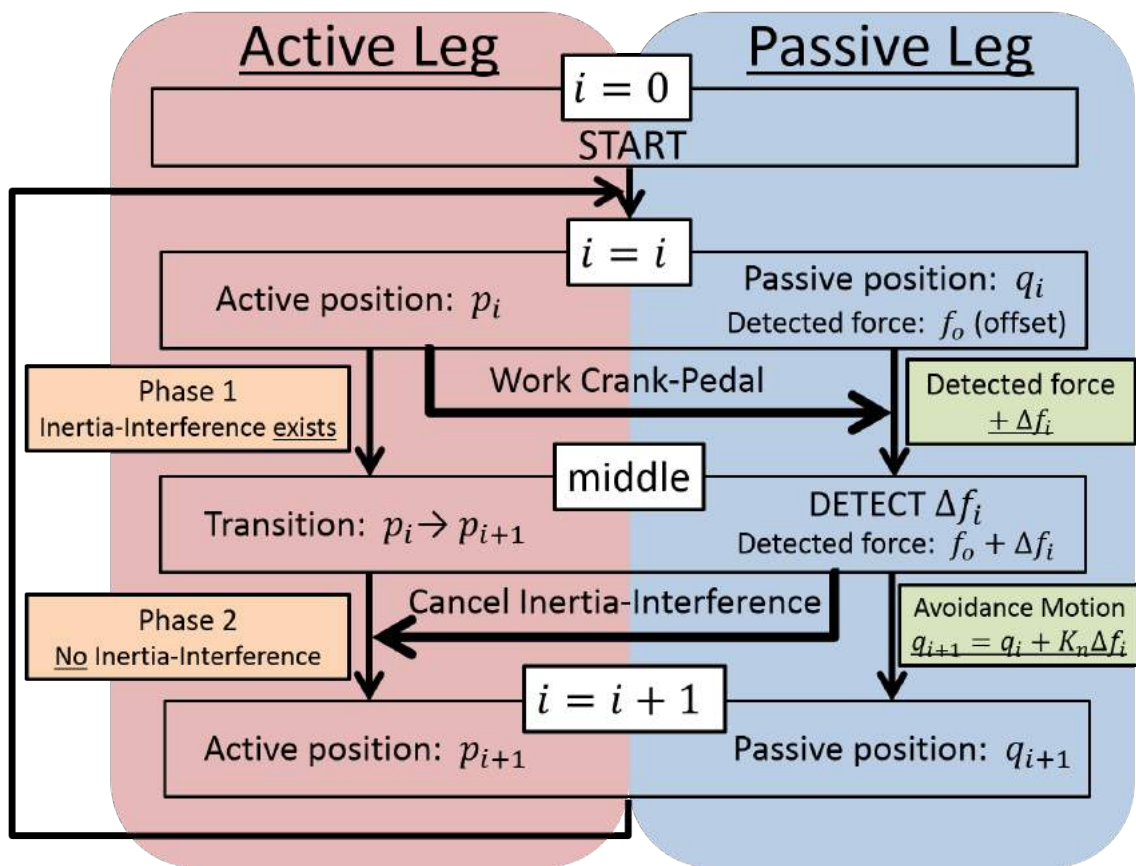


図 A.17: Process about Cooperative Controller of Active Leg and Passive Leg [126]

発表文献

筆頭著者論文

学術論文誌

1. 木村 航平, 溝花 弘登, 野沢 峻一, 岡田 慧, 稲葉 雅幸. 制御実行中にオンラインで PID ゲインをチューニングする学習制御に基づく等身大ヒューマノイドロボットの平行二輪電動スクーター上での調速行動. 日本機械学会論文集, Vol. 84 No. 864 p. 18-00032, 2018.
2. 木村 航平, 野沢 峻一, 垣内 洋平, 岡田 慧, 稲葉 雅幸. 能動機能と受動機能の切替制御システムを備えたヒューマノイドの双腕双脚による三輪車操作行動. 日本ロボット学会誌, Vol. 34 No. 7 pp. 468-477, 2016.
3. 木村 航平, 浅野 悠紀, 上月 豊隆, 白井 拓磨, 岡田 慧, 稲葉 雅幸. 能動脚と受動脚の協調システムに基づく腿駆動ヒューマノイドによるペダリング操作の実現. 計測自動制御学会論文集, Vol. 52 No. 8 pp. 428-436, 2016.

国際会議

4. Kohei Kimura, Shunichi Nozawa, Hiroto Mizohana, Kei Okada, Masayuki Inaba. Riding and speed governing for parallel two-wheeled scooter based on sequential online learning control by humanoid robot. In *Proceedings of The 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2018)*, (accepted), 2018.
5. Kohei Kimura, Shunichi Nozawa, Yohei Kakiuchi, Kei Okada, Masayuki Inaba. Tricycle manipulation strategy for humanoid robot based on active and passive manipulators control. In *Proceedings of The 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2016)*, pp. 5797-5804, 2016.

国内学術講演会

6. 木村 航平, 溝花 弘登, 野沢 峻一, 岡田 慧, 稲葉 雅幸. 制御実行中にオンラインで PID ゲインをチューニングする学習制御に基づく等身大ヒューマノイドロボットの平行二輪電動スクーター上での調速行動. 第 23 回ロボティクスシンポジウム予稿集, pp. 16-19, 2018.
7. 木村 航平, 野沢 峻一, 垣内 洋平, 岡田 慧, 稲葉 雅幸. 能動機能と受動機能を備えた双腕双脚の制御に基づくヒューマノイドによる三輪車操作と座り補正動作の実現. 第 33 回日本ロボット学会学術講演会講演論文集, 1G3-04, 2015.
8. 木村 航平, 浅野 悠紀, 上月 豊隆, 白井 拓磨, 岡田 慧, 稲葉 雅幸. 能動脚と受動脚の交互協調によるヒューマノイドのペダリング操作と走行の実現. 第 15 回 SICE システムインテグレーション部門講演会講演概要集, pp. 1194-1197, 2014.

9. 木村 航平, 浅野 悠紀, 白井 拓磨, 上月 豊隆, 茂木 陽太郎, 中島 慎介, 岡田 慧, 稲葉 雅幸. 腱駆動足首の柔軟弾性特性を利用したヒューマノイドのペダリング適応動作の実現. 日本機械学会ロボティクス・メカトロニクス講演会'14 講演論文集, 3P1-F05, 2014.

受賞

10. 木村 航平, 浅野 悠紀, 上月 豊隆, 白井 拓磨, 岡田 慧, 稲葉 雅幸. 能動脚と受動脚の交互協調によるヒューマノイドのペダリング操作と走行の実現. 計測自動制御学会システムインテグレーション部門講演会 *SI2014* 優秀講演賞, 第 15 回 SICE システムインテグレーション部門講演会, 2014.12.16.

共著論文

学術論文誌

1. Krishneel Chand Chaudhary, Kentaro Wada, Xiangyu Chen, Kohei Kimura, Kei Okada, Masayuki Inaba. Learning to segment generic handheld objects using class-agnostic deep comparison and segmentation network. *IEEE Robotics and Automation Letters*, (accepted), 2018.

国際会議

2. Xiangyu Chen, Kohei Kimura, Hiroto Mizohana, Moju Zhao, Fan Shi, Krishneel Chaudhary, Wesley P. Chan, Shunichi Nozawa, Yohei Kakiuchi, Kei Okada, Masayuki Inaba. Development of task-oriented high power field robot platform with humanoid upper body and mobile wheeled base. In *Proceedings of The 2016 IEEE/SICE International Symposium on System Integration*, pp. 349-354, 2016.
3. Yuki Asano, Soichi Ookubo, Toyotaka Kozuki, Takuma Shirai, Kohei Kimura, Shunichi Nozawa, Yohei Kakiuchi, Kei Okada, Masayuki Inaba. Spine balancing strategy using muscle ZMP on musculoskeletal humanoid Kenshiro. In *Proceedings of The 2015 International Symposium on Robotics Research*, pp. 1-16, 2015.
4. Yuki Asano, Toyotaka Kozuki, Soichi Ookubo, Koji Kawasaki, Takuma Shirai, Kohei Kimura, Kei Okada, Masayuki Inaba. A sensor-driver integrated muscle module with high-tension measurability and flexibility for tendon-driven robots. In *Proceedings of The 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5960-5965, 2015.

国内学術講演会

5. 溝花 弘登, Chan Wesley, 木村 航平, 陳 相羽, 岡田 慧, 稲葉 雅幸. 移動台車型ヒューマノイドによる屋外環境での工具取得操作行動における環境と自己の陰影を考慮する認識行動. 第17回 SICE システムインテグレーション部門講演会講演概要集, pp. 2611-2616, 2016.
6. 浅野 悠紀, 上月 豊隆, 川崎 宏治, 茂木 陽太郎, 趙 漢居, 白井 拓磨, 大久保 壮一, 木村 航平, 矢口 裕明, 垣内 洋平, 岡田 慧, 稲葉 雅幸. 反射的振る舞いに対する自動車衝突試験のための筋骨格ヒューマノイドによる運転と衝突反射行動の実現. 第32回日本ロボット学会学術講演会講演論文集, 2D1-01, 2014.

参考文献

- [1] 山下忠. 機械で操作される人工の指とそのマテリアルズハンドリングへの応用. 計測と制御, Vol. 3, No. 6, pp. 429–439, 1964.
- [2] 井上博允. 人工の手の計算機制御. 日本機械学会誌, Vol. 73, No. 618, pp. 946–954, 1970.
- [3] 内山勝. 人工の手の運動制御に関する研究 (第1報, 特異点を考慮した協調運動の計算). 日本機械学会論文集 (C編), Vol. 45, No. 391, pp. 314–322, 1979.
- [4] 内山勝. 人工の手の運動制御に関する研究 (第2報, 数学モデルによる協調運動の制御). 日本機械学会論文集 (C編), Vol. 45, No. 391, pp. 323–335, 1979.
- [5] 内山勝. 人工の手の運動制御に関する研究 (第3報, 動的視覚とビジュアルフィードバック). 日本機械学会論文集 (C編), Vol. 45, No. 391, pp. 336–345, 1979.
- [6] 自動車組立用にスポット溶接ロボットの大量導入始まる. http://rraj.rsj-web.org/back_wp/wp-content/uploads/x12_01.pdf.
- [7] 松元逸郎, 久良修郭, 仁尾理, 藤井肇, 鈴木健生. 産業用ロボット Motoman-L. 安川電機, 通巻第161号, Vol. 42, No. 3, pp. 184–193, 1978.
- [8] 下山勲. 竹馬型2足歩行ロボットの動的歩行. 日本機械学会論文集 (C編), Vol. 48, No. 433, pp. 1445–1455, 1982.
- [9] 梶田秀司, 谷和男. 線形倒立振りモードを規範とする凹凸路面上の動的2足歩行制御. 計測自動制御学会論文集, Vol. 31, No. 10, pp. 1705–1714, 1995.
- [10] 山口仁一, 高西淳夫, 加藤一郎. 路面形状に偏差のある環境における2足歩行制御—未知の凹凸路面に対する適応歩行の実現—. 日本ロボット学会誌, Vol. 13, No. 7, pp. 1030–1037, 1995.
- [11] 山口仁一, 木下昇, 高西淳夫, 加藤一郎. 路面形状に偏差のある環境に対する適応能力を持つ2足歩行ロボットの開発. 日本ロボット学会誌, Vol. 14, No. 4, pp. 546–559, 1996.
- [12] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki, and K. Fujimura. The intelligent ASIMO: System overview and integration. In *Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2478–2483, 2002.
- [13] 赤地一彦, 五十棲隆勝, 平田勝, 太田成彦, 石崎雅一. ヒューマノイドロボット・HRP-2の開発～人と一緒に働くロボットの設計・製作について～. 川田技報, Vol. 16, pp. 20–25, 2004.
- [14] 五十棲隆勝, 赤地一彦, 平田勝, 金子健二, 梶田秀司, 比留川博久. ヒューマノイドロボット HRP-2の開発. 日本ロボット学会誌, Vol. 22, No. 8, pp. 1004–1012, 2004.
- [15] Y. Ogura, H. Aikawa, K. Shimomura, H. Kondo, A. Morishima, H. o. Lim, and A. Takanishi. Development of a new humanoid robot WABIAN-2. In *Proceedings of The 2006 IEEE International Conference on Robotics and Automation*, pp. 76–81, 2006.
- [16] K. Kojima, T. Karasawa, T. Kozuki, E. Kuroiwa, S. Yukizaki, S. Iwaishi, T. Ishikawa, R. Koyama, S. Noda, F. Sugai, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Development of Life-Sized High-Power Humanoid Robot JAXON for Real-World Use. In *Proceedings of the 2015 IEEE-RAS International Conference on Humanoid Robots*, pp. 373–380, 2015.
- [17] 小島邦生, 唐澤達史, 上月豊隆, 黒岩英則, 柚木崎創, 岩石智志, 石川達矢, 小山遼, 野田晋太郎, 植田亮平, 菅井文仁, 野沢峻一, 垣内洋平, 岡田慧, 稲葉雅幸. 高速大出力ヒューマノイドの研究用プラットフォーム JAXON の開発. 日本ロボット学会誌, Vol. 34, No. 7, pp. 458–467, 2016.
- [18] I. W. Park, J. Y. Kim, J. Lee, and J. H. Oh. Mechanical design of humanoid robot platform KHR-3 (KAIST Humanoid Robot 3: HUBO). In *Proceedings of the 2005 IEEE-RAS International Conference on Humanoid Robots*, pp. 321–326, 2005.

- [19] T. Jung, J. Lim, H. Bae, K. K. Lee, H. M. Joe, and J. H. Oh. Development of the Humanoid Disaster Response Platform DRC-HUBO+. *IEEE Transactions on Robotics*, Vol. 34, No. 1, pp. 1–17, 2018.
- [20] Atlas. <https://www.bostondynamics.com/atlas>.
- [21] X. He, T. Ogura, A. Satou, and O. Hasegawa. Developmental Word Acquisition and Grammar Learning by Humanoid Robots Through a Self-Organizing Incremental Neural Network. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 37, No. 5, pp. 1357–1372, 2007.
- [22] 巻瀧有哉, 申富饒, 長谷川修. 実世界における一般問題解決システムの提案とそのヒューマノイドロボットへの実装. *電子情報通信学会論文誌 D*, Vol. J93-D, No. 6, pp. 960–977, 2010.
- [23] 高野渉. ヒューマノイドロボットの知能の設計: 身体運動の記号化から自然言語への展開. *人工知能学会誌*, Vol. 27, No. 6, pp. 587–593, 2012.
- [24] 野田邦昭, 有江浩明, 菅佑樹, 尾形哲也. Deep neural network を用いたヒューマノイドロボットによる物体操作行動の記憶学習と行動生成. 第 27 回 人工知能学会全国大会, 2013, pp. 2G4–OS–19a–2, 2013.
- [25] K. Wada, M. Murooka, K. Okada, and M. Inaba. 3D object segmentation for shelf bin picking by humanoid with deep learning and occupancy voxel grid map. In *Proceedings of the 2016 IEEE-RAS International Conference on Humanoid Robots*, pp. 1149–1154, 2016.
- [26] 室岡雅樹, 二井谷勇佑, 和田健太郎, 野沢峻一, 垣内洋平, 岡田慧, 稲葉雅幸. ロボットマニピュレーションにおける画像内物体の深層学習による運動予測. 第 31 回 人工知能学会全国大会, 2017, pp. 2B1–1, 2017.
- [27] Pepper. <https://www.softbank.jp/robot/>.
- [28] T. Mar, V. Tikhonoff, G. Metta, and L. Natale. Self-supervised learning of grasp dependent tool affordances on the iCub Humanoid robot. In *Proceedings of The 2015 IEEE International Conference on Robotics and Automation*, pp. 3200–3206, 2015.
- [29] Y. Ohara, M. Murooka, R. Ueda, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Configurable autonomy applicable to humanoid manipulation in unstructured and communication-limited environment. In *Proceedings of the 2015 IEEE-RAS International Conference on Humanoid Robots*, pp. 373–380, 2015.
- [30] S. Komatsu, Y. Kakiuchi, S. Nozawa, Y. Kojio, F. Sugai, K. Okada, and M. Inaba. Tool force adaptation in soil-digging task for humanoid robot. In *Proceedings of the 2017 IEEE-RAS International Conference on Humanoid Robots*, pp. 378–383, 2017.
- [31] H. Hasunuma, M. Kobayashi, H. Moriyama, T. Itoko, Y. Yanagihara, T. Ueno, K. Ohya, and K. Yokoi. A tele-operated humanoid robot drives a lift truck. In *Proceedings of The 2002 IEEE International Conference on Robotics and Automation*, pp. 2246–2252, 2002.
- [32] H. Hasunuma, K. Nakashima, M. Kobayashi, F. Mifune, Y. Yanagihara, T. Ueno, K. Ohya, and K. Yokoi. A tele-operated humanoid robot drives a backhoe. In *Proceedings of The 2003 IEEE International Conference on Robotics and Automation*, pp. 2998–3004 vol.3, 2003.
- [33] K. Yokoi, K. Nakashima, M. Kobayashi, H. Mihune, H. Hasunuma, Y. Yanagihara, T. Ueno, T. Gokuyuu, and K. Endou. A tele-operated humanoid robot drives a backhoe in the open air. In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1117–1122, 2003.
- [34] DARPA Robotics Challenge (DRC) (Archived). <https://www.darpa.mil/program/darpa-robotics-challenge>.

- [35] C. Rasmussen, K. Sohn, Q. Wang, and P. Oh. Perception and control strategies for driving utility vehicles with a humanoid robot. In *Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 973–980, 2014.
- [36] K. Knoedler, V. Dimitrov, D. Conn, M. A. Gennert, and T. Padr. Towards supervisory control of humanoid robots for driving vehicles during disaster response missions. In *2015 IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*, pp. 1–6, 2015.
- [37] I. Kumagai, R. Terasawa, S. Noda, R. Ueda, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Achievement of recognition guided teleoperation driving system for humanoid robots with vehicle path estimation. In *Proceedings of the 2015 IEEE-RAS International Conference on Humanoid Robots*, pp. 670–675, 2015.
- [38] H. Jeong, J. Oh, M. Kim, K. Joo, I. S. Kweon, and J. H. Oh. Control strategies for a humanoid robot to drive and then egress a utility vehicle for remote approach. In *Proceedings of the 2015 IEEE-RAS International Conference on Humanoid Robots*, pp. 811–816, 2015.
- [39] J. Rojas, W. Newman, H. Jie, and Q. Liu. A steering wheel manipulation scheme by an anthropomorphic humanoid robot in a constrained vehicle environment. In *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1566–1571, 2015.
- [40] A. Paolillo, P. Gergondet, A. Cherubini, M. Vendittelli, and A. Kheddar. Autonomous car driving by a humanoid robot. *Journal of Field Robotics*, Vol. 35, No. 2, pp. 169–186, 2017.
- [41] T. Yoshikai, Y. Nakanishi, I. Mizuuchi, and M. Inaba. Pedaling Motion of a Cycle by Musculoskeletal Humanoid with Adapting Ability Based on an Evaluation of the Muscle Loads. In *Proceedings of the 9th International Conference on Intelligent Autonomous Systems (IAS)*, pp. 767–775, 2006.
- [42] Y. Nakanishi, I. Mizuuchi, T. Yoshikai, T. Inamura, and M. Inaba. Pedaling by a redundant musculoskeletal humanoid robot. In *Proceedings of the 2005 IEEE-RAS International Conference on Humanoid Robots*, pp. 68–73, 2005.
- [43] T. Oku, K. Inoue, T. T. H. Pham, K. Tominaga, D. Maeda, M. Uemura, H. Hirai, and F. Miyazaki. Analysis of muscle coordination in human pedaling and implementation with a musculoskeletal robot. In *Proceedings of the 2012 IEEE-RAS International Conference on Humanoid Robots*, pp. 606–611, 2012.
- [44] A. Gams, J. v. d. Kieboom, M. Vespignani, L. Guyot, A. Ude, and A. Ijspeert. Rich periodic motor skills on humanoid robots: Riding the pedal racer. In *Proceedings of The 2014 IEEE International Conference on Robotics and Automation*, pp. 2326–2332, 2014.
- [45] 日本工業規格. JIS B 0186:2003 移動ロボット-用語. <http://kikakurui.com/b0/B0186-2003-01.html>.
- [46] 涌井伸二, 美多勉. 連成振動による慣性項同定精度の劣化とその改善. *日本ロボット学会誌*, Vol. 7, No. 3, pp. 195–198, 1989.
- [47] 大明準治. ロボット運動制御のための動力学モデリングと物理パラメータ推定技術. *東芝レビュー*, Vol. 67, No. 7, pp. 33–37, 2012.
- [48] 吉川恒夫. マニピュレータの力制御. *計測と制御*, Vol. 30, No. 5, pp. 383–388, 1991.
- [49] 小菅一弘. 力制御法の分類と制御システムの設計法. *日本ロボット学会誌*, Vol. 9, No. 6, pp. 751–758, 1991.
- [50] 武居直行, 野畑茂広, 藤本英雄. 安定した接触を実現するアドミッタンス制御手法. *日本ロボット学会誌*, Vol. 26, No. 6, pp. 635–642, 2008.

- [51] 河原塚健人, 牧野将吾, 川村将矢, 浅野悠紀, 岡田慧, 稲葉雅幸. 筋骨格ヒューマノイドにおける視覚を利用した関節-筋空間マップの逐次的再学習. 第35回日本ロボット学会学術講演会講演論文集, pp. 2L1-01, 2017.
- [52] 木村航平, 野沢峻一, 垣内洋平, 岡田慧, 稲葉雅幸. 能動機能と受動機能の切替制御システムを備えたヒューマノイドの双腕双脚による三輪車操作行動. 日本ロボット学会誌, Vol. 34, No. 7, pp. 468-477, 2016.
- [53] K. Kimura, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Tricycle Manipulation Strategy for Humanoid Robot Based on Active and Passive Manipulators Control. In *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5797-5804, 2016.
- [54] 木村航平, 野沢峻一, 垣内洋平, 岡田慧, 稲葉雅幸. 能動機能と受動機能を備えた双腕双脚の制御に基づくヒューマノイドによる三輪車操作と座り補正動作の実現. 第33回日本ロボット学会学術講演会講演論文集, pp. 1G3-04, 2015.
- [55] S. Kajita, M. Morisawa, K. Miura, S. Nakaoka, K. Harada, K. Kaneko, F. Kanehiro, and K. Yokoi. Biped walking stabilization based on linear inverted pendulum tracking. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4489-4496, 2010.
- [56] S. H. Hyon, J. G. Hale, and G. Cheng. Full-Body Compliant Human-Humanoid Interaction: Balancing in the Presence of Unknown External Forces. *IEEE Transactions on Robotics*, Vol. 23, No. 5, pp. 884-898, 2007.
- [57] J. Pratt, J. Carff, S. Drakunov, and A. Goswami. Capture Point: A Step toward Humanoid Push Recovery. In *Proceedings of the 2006 IEEE-RAS International Conference on Humanoid Robots*, pp. 200-207, 2006.
- [58] Vivo Market 株式会社. V-BOARD. <http://vivo-market.com/product/>.
- [59] 木村航平, 溝花弘登, 野沢峻一, 岡田慧, 稲葉雅幸. 制御実行中にオンラインでPIDゲインをチューニングする学習制御に基づく等身大ヒューマノイドロボットの平行二輪電動スクーター上での調速行動. 第23回ロボティクスシンポジウム予稿集, pp. 16-19, 2018.
- [60] M. Vukobratović and J. Stepanenko. On the stability of anthropomorphic systems. *Mathematical Biosciences*, Vol. 15, No. 1-2, pp. 1-37, 1972.
- [61] M. W. M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotics and Automation*, Vol. 17, No. 3, pp. 229-241, 2001.
- [62] S. Thrun. Robotic mapping: A survey. *Exploring artificial intelligence in the new millennium*, Vol. 1, pp. 1-35, 2002.
- [63] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse. MonoSLAM: Real-Time Single Camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 6, pp. 1052-1067, 2007.
- [64] I. Kumagai, R. Ueda, F. Sugai, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Achievement of localization system for humanoid robots with virtual horizontal scan relative to improved odometry fusing internal sensors and visual information. In *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 666-673, 2016.
- [65] I. Kumagai, F. Sugai, S. Nozawa, Y. Kakiuchi, K. Okada, M. Inaba, and F. Kanehiro. Complementary integration framework for localization and recognition of a humanoid robot based on task-oriented frequency and accuracy requirements. In *Proceedings of the 2017 IEEE-RAS International Conference on Humanoid Robots*, pp. 683-688, 2017.

- [66] K. Hsiao, S. Chitta, M. Ciocarlie, and E. G. Jones. Contact-reactive grasping of objects with partial shape information. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1228–1235, 2010.
- [67] N. Sommer, M. Li, and A. Billard. Bimanual compliant tactile exploration for grasping unknown objects. In *Proceedings of The 2014 IEEE International Conference on Robotics and Automation*, pp. 6400–6407, 2014.
- [68] M. Murooka, R. Ueda, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Planning and Execution of Groping Behavior for Contact Sensor based Manipulation in an Unknown Environment. In *Proceedings of The 2016 IEEE International Conference on Robotics and Automation*, pp. 3955–3962, 2016.
- [69] C. C. Wong, C. T. Cheng, K. H. Huang, Y. T. Yang, H. M. Chan, and C. S. Yin. Obstacle Avoidance Design for Humanoid Robot Based on Four Infrared Sensors. *Tamkang Journal of Science and Engineering*, Vol. 12, No. 3, pp. 249–258, 2009.
- [70] I. Kumagai, S. Noda, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Whole body joint load reduction control for high-load tasks of humanoid robot through adapting joint torque limitation based on online joint temperature estimation. In *Proceedings of the 2014 IEEE-RAS International Conference on Humanoid Robots*, pp. 463–468, 2014.
- [71] Y. Ito, S. Nozawa, J. Urata, T. Nakaoka, K. Kobayashi, Y. Nakanishi, K. Okada, and M. Inaba. Development and verification of life-size humanoid with high-output actuation system. In *Proceedings of The 2014 IEEE International Conference on Robotics and Automation*, pp. 3433–3438, 2014.
- [72] X. Chen, K. Kimura, H. Mizohana, M. Zhao, F. Shi, K. Chaudhary, W. P. Chan, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Development of task-oriented high power field robot platform with humanoid upper body and mobile wheeled base. In *2016 IEEE/SICE International Symposium on System Integration (SII)*, pp. 349–354, 2016.
- [73] Y. Hosoda, S. Egawa, J. Tamamoto, K. Yamamoto, R. Nakamura, and M. Togami. Basic Design of Human-Symbiotic Robot EMIEW. In *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5079–5084, 2006.
- [74] K. Hirai. Current and future perspective of Honda humanoid robot. In *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 500–508, 1997.
- [75] H. Yaguchi, K. Sasabuchi, W. P. Chan, K. Nagahama, T. Saiki, Y. Shiigi, and M. Inaba. A design of 4-legged semi humanoid robot aero for disaster response task. In *Proceedings of the 2015 IEEE-RAS International Conference on Humanoid Robots*, pp. 61–66, 2015.
- [76] 株式会社村田製作所. ムラタセイコちゃん. <https://www.murata.com/ja-jp/about/mboyngirl/mgirl>.
- [77] ヤマハ発動機株式会社. MOTOBOT. <https://global.yamaha-motor.com/jp/showroom/motobot/>.
- [78] 日本工業規格. JIS B 0185:2002 知能ロボット-用語. <http://kikakurui.com/b0/B0185-2002-01.html>.
- [79] move_base - ROS Wiki. http://wiki.ros.org/move_base.
- [80] I. M. Ariffin, A. Baharuddin, A. C. Atien, and H. Yussof. Real-Time Obstacle Avoidance for Humanoid-Controlled Mobile Platform Navigation. *Procedia Computer Science*, Vol. 105, pp. 34–39, 2017.

- [81] H. Hjalmarsson. Iterative feedback tuning : an overview. *International Journal of Adaptive Control and Signal Processing*, Vol. 16, No. 5, pp. 373–395, 2002.
- [82] F. D. Bruyne, B. D. O. Anderson, M. Gevers, and N. Linard. Iterative controller optimization for nonlinear systems. In *Proceedings of the 36th IEEE Conference on Decision and Control*, Vol. 4, pp. 3749–3754, 1997.
- [83] 相馬将太郎, 金子修, 藤井隆雄. 一回の実験データに基づく制御器パラメータチューニングの新しいアプローチ : Fictitious Reference Iterative Tuning の提案. システム制御情報学会論文誌, Vol. 17, No. 12, pp. 528–536, 2004.
- [84] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi, and H. Hirukawa. Biped walking pattern generation by using preview control of zero-moment point. In *Proceedings of The 2003 IEEE International Conference on Robotics and Automation*, pp. 1620–1626, 2003.
- [85] S. Nozawa, E. Kuroiwa, K. Kojima, R. Ueda, M. Murooka, S. Noda, I. Kumagai, Y. Ohara, Y. Kakiuchi, K. Okada, and M. Inaba. Multi-layered real-time controllers for humanoid’s manipulation and locomotion tasks with emergency stop. In *Proceedings of the 2015 IEEE-RAS International Conference on Humanoid Robots*, pp. 381–388, 2015.
- [86] N. Ratliff, M. Zucker, J. A. Bagnell, and S. Srinivasa. CHOMP: Gradient optimization techniques for efficient motion planning. In *Proceedings of The 2009 IEEE International Conference on Robotics and Automation*, pp. 489–494, 2009.
- [87] M. Zucker, N. Ratliff, A. Dragan, M. Pivtoraiko, M. Klingensmith, C. Dellin, J. Bagnell, and S. Srinivasa. CHOMP: Covariant Hamiltonian Optimization for Motion Planning. *The International Journal of Robotics Research*, Vol. 32, No. 9-10, pp. 1164–1193, 2013.
- [88] M. Kalakrishnan, S. Chitta, E. Theodorou, P. Pastor, and S. Schaal. STOMP: Stochastic trajectory optimization for motion planning. In *Proceedings of The 2011 IEEE International Conference on Robotics and Automation*, pp. 4569–4574, 2011.
- [89] A. Bonarini, C. Caccia, A. Lazaric, and M. Restelli. Batch Reinforcement Learning for Controlling a Mobile Wheeled Pendulum Robot. In M. Bramer, editor, *IFIP International Conference on Artificial Intelligence in Theory and Practice*, pp. 151–160, Boston, MA, 2008. Springer US.
- [90] M. Riedmiller, T. Gabel, R. Hafner, and S. Lange. Reinforcement learning for robot soccer. *Autonomous Robots*, Vol. 27, No. 1, pp. 55–73, 2009.
- [91] J. Cunha, R. Serra, N. Lau, L. Lopes, and A. Neves. Batch Reinforcement Learning for Robotic Soccer Using the Q-Batch Update-Rule. *Journal of Intelligent & Robotic Systems*, Vol. 80, No. 3, pp. 385–399, 2015.
- [92] L. Bottou. Online learning and stochastic approximations. *On-Line Learning in Neural Networks*, pp. 9–42, 1998.
- [93] N. Qian. On the momentum term in gradient descent learning algorithms. *Neural Networks*, Vol. 12, No. 1, pp. 145–151, 1999.
- [94] J. Duchi, E. Hazan, and Y. Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, Vol. 12, pp. 2121–2159, 2011.
- [95] S. Nozawa, R. Ueda, Y. Kakiuchi, K. Okada, and M. Inaba. A full-body motion control method for a humanoid robot based on on-line estimation of the operational force of an object with an unknown weight. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2684–2691, 2010.

- [96] M. Murooka, S. Noda, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Manipulation strategy decision and execution based on strategy proving operation for carrying large and heavy objects. In *Proceedings of The 2014 IEEE International Conference on Robotics and Automation*, pp. 3425–3432, 2014.
- [97] S. Nozawa, S. Noda, M. Murooka, K. Okada, and M. Inaba. Online estimation of object-environment constraints for planning of humanoid motion on a movable object. In *Proceedings of The 2017 IEEE International Conference on Robotics and Automation*, pp. 1291–1298, 2017.
- [98] M. Murooka, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Feasibility evaluation of object manipulation by a humanoid robot based on recursive estimation of the object’s physical properties. In *Proceedings of The 2017 IEEE International Conference on Robotics and Automation*, pp. 4082–4089, 2017.
- [99] 武田常広, 北森俊行. 2次形式評価関数の重み行列の一決定法. 計測自動制御学会論文集, Vol. 17, No. 6, pp. 631–638, 1981.
- [100] M. R. Cutkosky. On Grasp Choice, Grasp Models, and the Design of Hands for Manufacturing Tasks. *IEEE Transactions on Robotics and Automation*, Vol. 5, No. 3, pp. 269–279, 1989.
- [101] N. Vahrenkamp, S. Wieland, P. Azad, D. Gonzalez, T. Asfour, and R. Dillmann. Visual Servoing for Humanoid Grasping and Manipulation Tasks. In *Proceedings of the 2008 IEEE-RAS International Conference on Humanoid Robots*, pp. 406–412, 2008.
- [102] K. Harada, S. Kajita, K. Kaneko, and H. Hirukawa. ZMP analysis for arm/leg coordination. In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 1, pp. 75–81, 2003.
- [103] L. Sentis, J. Park, and O. Khatib. Modeling and control of multi-contact centers of pressure and internal forces in humanoid robots. In *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 453–460, 2009.
- [104] 長阪憲一郎, 福島哲治, 下村秀樹. 接触拘束を考慮可能なマルチコンタクト対応スタビライザと一般化逆動力学による人型ロボットの全身制御. 第17回ロボティクスシンポジウム予稿集, pp. 134–141, 2012.
- [105] 相山康道. グラスプレス・マニピュレーションに関する研究. PhD thesis, 東京大学大学院工学系研究科情報工学専攻, 1995.
- [106] F. Kanehiro, T. Yoshimi, S. Kajita, M. Morisawa, K. Fujiwara, K. Harada, K. Kaneko, H. Hirukawa, and F. Tomita. Whole Body Locomotion Planning of Humanoid Robots based on a 3D Grid Map. In *Proceedings of The 2005 IEEE International Conference on Robotics and Automation*, pp. 1072–1078, 2005.
- [107] C. Park, J. H. Kyung, T. Y. Choi, H. M. Do, B. I. Kim, and S. H. Lee. Design of an industrial dual arm robot manipulator for a Human-Robot hybrid manufacturing. In *2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pp. 616–618, 2012.
- [108] S. Y. Shin, J. w. Lee, and C. Kim. Humanoid’s dual arm object manipulation based on virtual dynamics model. In *Proceedings of The 2012 IEEE International Conference on Robotics and Automation*, pp. 2599–2604, 2012.
- [109] 浅野悠紀, 上月豊隆, 川崎宏治, 茂木陽太郎, 趙漢居, 白井拓磨, 大久保壮一, 木村航平, 矢口裕明, 垣内洋平, 岡田慧, 稲葉雅幸. 反射的振る舞いに対する自動車衝突試験のための筋骨格ヒューマノイドによる運転と衝突反射行動の実現. 日本ロボット学会第32回学術講演会予稿集, 2014.

- [110] K. Okada, T. Ogura, A. Haneda, J. Fujimoto, F. Gravot, and M. Inaba. Humanoid motion generation system on HRP2-JSK for daily life environment. In *2005 IEEE International Conference Mechatronics and Automation*, Vol. 4, pp. 1772–1777, 2005.
- [111] C. Klein and C. Huang. Review of pseudoinverse control for use with kinematically redundant manipulators. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-13, No. 2, pp. 245–250, 1983.
- [112] O. Khatib. A Unified Approach for Motion and Force Control of Robot Manipulators: The Operational Space. *IEEE Journal of Robotics and Automation*, Vol. 3, No. 1, pp. 43–53, 1987.
- [113] 吉田和哉, 倉爪亮, 梅谷陽二. 宇宙ロボットの多腕強調制御. 日本ロボット学会誌, Vol. 6, No. 9, pp. 34–42, 1991.
- [114] J. Kerr and B. Roth. Analysis of Multifingered Hands. *The International Journal of Robotics Research*, Vol. 4, No. 4, pp. 3–17, 1986.
- [115] E. Nakano, S. Ozaki, T. Ishida, and I. Kato. Cooperational Control of the Anthropomorphic Manipulator 'MELARM'. In *4th International Symposium on Industrial Robots*, pp. 251–260, 1974.
- [116] 中村仁彦, 永井清, 吉川恒夫. 複数のロボット機構による協調的あやつりの力学. 日本ロボット学会誌, Vol. 4, No. 5, pp. 23–32, 1986.
- [117] Y. F. Zheng and J. Y. S. Luh. Optimal Load Distribution for Two Industrial Robots Handling a Single Object. In *Proceedings of The 1988 IEEE International Conference on Robotics and Automation*, pp. 344–349, 1988.
- [118] M. Uchiyama. A Unified Approach to Load Sharing, Motion Decomposing and Force Sensing of Dual Arm Robots. In *Robotics Research: The Fifth International Symposium*, pp. 225–232. MIT press, 1990.
- [119] Yoshikawa, T. and Zheng, X. Coordinated dynamic hybrid position/force control for multiple robot manipulators handling one constrained object. In *Proceedings of The 1990 IEEE International Conference on Robotics and Automation*, 1990.
- [120] 吉川恒夫, 鄭心知. 拘束運動下にある対象物の複数台ロボットによる協調ハイブリッド制御. システム制御情報学会論文誌, Vol. 3, No. 10, pp. 326–334, 1990.
- [121] 小菅一弘, 吉田英博, 福田敏男, 蟹谷清, 酒井勝, 針木和夫. インピーダンス制御に基づく双腕マニピュレータの協調制御. 日本ロボット学会誌, Vol. 13, No. 3, pp. 404–410, 1995.
- [122] Y. Nakanishi, I. Mizuuchi, T. Yoshikai, T. Inamura, and M. Inaba. Pedaling by a redundant musculoskeletal humanoid robot. In *Proceedings of the 2005 IEEE-RAS International Conference on Humanoid Robots*, pp. 68–73, 2005.
- [123] T. Oku, K. Inoue, P. T. T. Hang, K. Tominaga, D. Maeda, M. Uemura, H. Hirai, and F. Miyazaki. Analysis of muscle coordination in human pedaling and implementation with a musculoskeletal robot. In *Proceedings of the 2012 IEEE-RAS International Conference on Humanoid Robots*, pp. 606–611, 2012.
- [124] 渡邊高広, 昆陽雅司, 永谷直久, 田所諭. ペダリング運動を用いた人体下肢のインピーダンス調整機能の評価、第4報:タスクを付与したペダリング運動中のペダル踏力計測. 第13回計測自動制御学会システムインテグレーション部門講演会 2012, pp. 1925–1926, 2012.
- [125] M. Gföhler and P. Lugner. Dynamic simulation of FES-cycling: influence of individual parameters. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 12, No. 4, pp. 398–405, 2004.

- [126] 木村航平, 浅野悠紀, 上月豊隆, 白井拓磨, 岡田慧, 稲葉雅幸. 能動脚と受動脚の協調システムに基づく腱駆動ヒューマノイドによるペダリング操作の実現. 計測自動制御学会論文集, Vol. 52, No. 8, pp. 428–436, 2016.
- [127] Y. Ohara, M. Murooka, R. Ueda, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Configurable autonomy applicable to humanoid manipulation in unstructured and communication-limited environment. In *Proceedings of the 2015 IEEE-RAS International Conference on Humanoid Robots*, pp. 373–380, 2015.
- [128] S. McGill, Y. Seung-Joon, and D. D. Lee. Team THOR’s adaptive autonomy for disaster response humanoids. In *Proceedings of the 2015 IEEE-RAS International Conference on Humanoid Robots*, pp. 453–460, 2015.
- [129] 矢野倉伊織, 小原由羽, 石黒康裕, 室岡雅樹, 野沢峻一, 岡田慧, 稲葉雅幸. 等身大ヒューマノイドの反復的道具利用による模倣動作の獲得 – HRP-2 による行動単位模倣獲得システム –. 第 22 回ロボティクスシンポジウム予稿集, pp. 171–176, 2017.
- [130] 野田晋太郎. 実体設計評価機構を備えたロボットの身体行動創成支援システムの構成法. PhD thesis, 東京大学大学院情報理工学系研究科創造情報学専攻, 2018.
- [131] N. Hogan. Impedance control: An approach to manipulation, part 2-implementation. *Journal of Dynamic systems, measurements, and control*, Vol. 107, pp. 8–16, 1985.
- [132] N. Hogan. Stable execution of contact tasks using impedance control. *Proceedings of The 1987 IEEE International Conference on Robotics and Automation*, Vol. 4, pp. 1047–1054, 1987.
- [133] F. Zacharias, C. Borst, and G. Hirzinger. Capturing robot workspace structure: representing robot capabilities. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3229–3236, 2007.
- [134] S. Nozawa, I. Kumagai, Y. Kakiuchi, K. Okada, and M. Inaba. Humanoid full-body controller adapting constraints in structured objects through updating task-level reference force. In *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3417–3424, 2012.
- [135] S. Jung, T. C. Hsia, and R. G. Bonitz. Force tracking impedance control of robot manipulators under unknown environment. *IEEE Transactions on Control Systems Technology*, Vol. 12, No. 3, pp. 474–483, 2004.
- [136] 中村仁彦. 物体把握とあやつりのための情報検出. 計測と制御, Vol. 26, No. 2, pp. 117–121, 1987.
- [137] D. Honegger, P. Greisen, L. Meier, P. Tanskanen, and M. Pollefeys. Real-time Velocity Estimation Based on Optical Flow and Disparity Matching. In *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5177–5182, 2012.
- [138] Y. Asano, H. Mizoguchi, T. Kozuki, Y. Motegi, M. Osada, J. Urata, Y. Nakanishi, K. Okada, and M. Inaba. Lower Thigh Design of Detailed Musculoskeletal Humanoid “Kenshiro”. In *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4367–4372, 2012.
- [139] Y. Nakanishi, Y. Asano, T. Kozuki, H. Mizoguchi, Y. Motegi, M. Osada, T. Shirai, J. Urata, K. Okada, and M. Inaba. Design Concept of Detail Musculoskeletal Humanoid “Kenshiro” - Toward a real human body musculoskeletal simulator -. In *Proceedings of the 2012 IEEE-RAS International Conference on Humanoid Robots*, pp. 1–6, 2012.
- [140] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng. ROS: an open-source Robot Operating System. In *ICRA Workshop on Open Source Software*, 2009.

- [141] 岡田慧. ROS (ロボット・オペレーティング・システム) . 日本ロボット学会誌, Vol. 30, No. 9, pp. 830–835, 2012.
- [142] M. Vukobratović, B. Borovac, D. Surla, and D. Stoki. Biped Locomotion: Dynamics, Stability, Control and Application. In *Vol. 7 of Scientific Fundamentals of Robotics*. Springer-Verlag, 1990.
- [143] K. Nishiwaki and S. Kagami. High Frequency Walking Pattern Generation based on Preview Control of ZMP. In *Proceedings of The 2006 IEEE International Conference on Robotics and Automation*, pp. 2667–2672, 2006.
- [144] T. Sugihara. Standing Stabilizability and Stepping Maneuver in Planar Bipedalism based on the Best COM-ZMP Regulator. In *Proceedings of The 2009 IEEE International Conference on Robotics and Automation*, pp. 1966–1971, 2009.
- [145] K. Harada, S. Kajita, K. Kaneko, and H. Hirukawa. ZMP analysis for arm/leg coordination. In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 75–81, 2003.
- [146] 原田研介, 梶田秀司, 金広文男, 藤原清司, 金子健二, 横井一仁, 比留川博久. ヒューマノイドロボットの脚腕協調における ZMP 解析. 日本ロボット学会誌, Vol. 22, No. 1, pp. 28–36, 2004.
- [147] M. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotics and Automation*, Vol. 17, No. 3, pp. 229–241, 2001.
- [148] Sabe, K. and Fukuchi, M. and Gutmann, J.S. and Ohashi, T. and Kawamoto, K. and Yoshigahara, T. Obstacle avoidance and path planning for humanoid robots using stereo vision. In *Proceedings of The 2004 IEEE International Conference on Robotics and Automation*, 2004.
- [149] A. Davison, I. Reid, N. Molton, and O. Stasse. MonoSLAM: Real-Time Single Camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 6, pp. 1052–1067, 2007.
- [150] Jingang Yi, Junjie Zhang, Dezhen Song, and Jayasuriya, S. IMU-based localization and slip estimation for skid-steered mobile robots. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2845–2850, 2007.
- [151] 梶田秀司, 森澤光晴, 三浦郁奈子, 中岡慎一郎, 金子健二, 金広文男, 横井一仁. 線形倒立振り子トラッキング制御による不整地 2 足歩行制御. 日本ロボット学会第 28 回学術講演会予稿集, pp. 2D2–8, 2010.
- [152] Q. Zhu. Hidden Markov model for dynamic obstacle avoidance of mobile robot navigation. *IEEE Transactions on Robotics and Automation*, Vol. 7, No. 3, pp. 390–397, 1991.
- [153] Okada, K. and Ogura, T. and Haneda, A. and Inaba, M. Autonomous 3D walking system for a humanoid robot based on visual step recognition and 3D foot step planner. In *Proceedings of The 2005 IEEE International Conference on Robotics and Automation*, 2005.
- [154] S. Jia, J. Sheng, D. Chugo, and K. Takase. Obstacle Recognition for a Mobile Robot in Indoor Environments using RFID and Stereo Vision. In *Proceedings of The 2005 IEEE International Conference on Mechatronics and Automation*, pp. 2789–2794, 2007.
- [155] MultiSense SL - Carnegie Robotics. <http://carnegierobotics.com/multisense-sl/>.
- [156] Y. Hwang and N. Ahuja. A potential field approach to path planning. *IEEE Transactions on Robotics and Automation*, Vol. 8, No. 1, pp. 23–32, 1992.

- [157] J. Mainprice, E. Sisbot, L. Jaillet, J. Cortes, R. Alami, and T. Simeon. Planning human-aware motions using a sampling-based costmap planner. In *Proceedings of The 2011 IEEE International Conference on Robotics and Automation*, pp. 5012–5017, 2011.
- [158] D. Fox, W. Burgard, and S. Thrun. Controlling synchro-drive robots with the dynamic window approach to collision avoidance. In *Proceedings of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 3, pp. 1280–1287, 1996.
- [159] 三平満司. 厳密な線形化とそのけん引車両の軌道制御への応用. 計測と制御, Vol. 31, No. 8, pp. 851–858, 1992.
- [160] H. Wang, Y. F. Zheng, Y. Jun, and P. Oh. DRC-hubo walking on rough terrains. In *2014 IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*, pp. 1–6, 2014.
- [161] C. Iverach-Brereton, J. Baltes, J. Anderson, A. Winton, and D. Carrier. Gait design for an ice skating humanoid robot. *Journal of Robotics and Autonomous Systems*, Vol. 62, No. 3, pp. 306–318, 2014.
- [162] N. Takasugi, K. Kojima, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba. Real-time skating motion control of humanoid robots for acceleration and balancing. In *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1356–1363, 2016.
- [163] N. Takasugi, K. Kojima, S. Nozawa, K. Okada, and M. Inaba. 3D walking and skating motion generation using divergent component of motion and gauss pseudospectral method. In *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5003–5010, 2017.
- [164] N. Ando, T. Suehiro, K. Kitagaki, T. Kotoku, and W.-K. Yoon. RT-middleware: distributed component middleware for RT (robot technology). In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3933–3938, 2005.
- [165] K. Fujimoto and T. Sugie. Iterative learning control of Hamiltonian systems: I/O based optimal control approach. *IEEE Transactions on Automatic Control*, Vol. 48, No. 10, pp. 1756–1761, 2003.
- [166] K. Hamamoto and T. Sugie. Iterative learning control for robot manipulators using the finite dimensional input subspace. *IEEE Transactions on Robotics and Automation*, Vol. 18, No. 4, pp. 632–635, 2002.
- [167] J. J. Kuffner, K. Nishiwaki, S. Kagami, M. Inaba, and H. Inoue. Footstep planning among obstacles for biped robots. In *Proceedings of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 500–505, 2001.
- [168] R. Deits and R. Tedrake. Footstep planning on uneven terrain with mixed-integer convex optimization. In *Proceedings of the 2014 IEEE-RAS International Conference on Humanoid Robots*, pp. 279–286, 2014.
- [169] K. Yamamoto. *Study on Adaptive Contact Force Distribution of Humanoid Robots - Force Distribution in Toe-Thenar Mechanism and Impedance Distribution between Legs -*. PhD thesis, Department of Intelligent Mechano-Informatics, Graduate School of University of Tokyo, 2009.
- [170] 山本江, 中村仁彦. 両脚床反力比に基づいたインピーダンス調整によるヒューマノイドの足裏路面適応. 第 26 回日本ロボット学会学術講演会講演論文集, pp. 2J1–05, 2008.
- [171] 野波健蔵, 水野毅, 足立修一, 池田雅夫, 大須賀公一, 大日方五郎, 木田隆, 永井正夫, 平田光男, 松野文俊. 制御の事典. 朝倉書店, 2015.
- [172] K. J. Åström and T. Hägglund. *PID Controllers: Theory, Design, and Tuning*. Instrument Society of America, 2 edition, 1995.

- [173] R. Hanus, M. Kinnaert, and J. L. Henrotte. Conditioning technique, a general anti-windup and bumpless transfer method. *Automatica*, Vol. 23, No. 6, pp. 729–739, 1987.
- [174] Y. Peng, D. Vrancic, and R. Hanus. Anti-windup, bumpless, and conditioned transfer techniques for PID controllers. *IEEE Control Systems Magazine*, Vol. 16, No. 4, pp. 48–57, 1996.
- [175] 鈴木文泰, 堀洋一. 飽和状態オブザーバを用いたアンチwindアップ制御. 電気学会論文誌 D, Vol. 120, No. 1, pp. 120–125, 2000.
- [176] A. Zheng, M. V. Kothare, and M. Morari. Anti-windup design for internal model control. *International Journal of Control*, Vol. 60, No. 5, pp. 1015–1024, 1994.
- [177] E. G. Gilbert, I. Kolmanovskiy, and K. T. Tan. Discrete-time reference governors and the nonlinear control of systems with state and control constraints. *International Journal of Robust and Nonlinear Control*, Vol. 5, No. 5, pp. 487–504, 1995.
- [178] 杉江俊治, 福井隆之, 鈴木宙見. 不確かさを考慮した拘束システムの目標値整形とその実験検証. システム制御情報学会論文誌, Vol. 17, No. 2, pp. 73–80, 2004.
- [179] 杉原知道. 目標値整形による簡単かつ安全な積分補償を導入したモータ制御. 日本ロボット学会誌, Vol. 27, No. 8, pp. 910–916, 2009.
- [180] A. J. Baerveldt and R. Klang. A low-cost and low-weight attitude estimation system for an autonomous helicopter. In *Proceedings of IEEE International Conference on Intelligent Engineering Systems*, pp. 391–395, 1997.
- [181] R. Mahony, T. Hamel, and J. M. Pflimlin. Nonlinear Complementary Filters on the Special Orthogonal Group. *IEEE Transactions on Automatic Control*, Vol. 53, No. 5, pp. 1203–1218, 2008.
- [182] E. J. Lefferts, F. L. Markley, and M. D. Shuster. Kalman filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, pp. 417–429, 1982.
- [183] Stereolabs Inc. ZED Stereo Camera. <https://www.stereolabs.com/>.
- [184] E. G. Gilbert and K. T. Tan. Linear systems with state and control constraints: the theory and application of maximal output admissible sets. *IEEE Transactions on Automatic Control*, Vol. 36, No. 9, pp. 1008–1020, 1991.
- [185] K. Hirata and M. Fujita. Set of admissible reference signals and control of systems with state and control constraints. In *Proceedings of the 38th IEEE Conference on Decision and Control*, Vol. 2, pp. 1427–1432, 1999.
- [186] 平田研二. あるクラスの非線形システムに対する最大出力許容集合の構成手順. 計測自動制御学会論文誌, Vol. 41, No. 6, pp. 544–546, 2005.
- [187] 太田有三, 関雄樹, 郭昊, 増淵泉. 最大出力許容集合の内側近似を用いた拘束系の制御. システム制御情報学会論文誌, Vol. 24, No. 10, pp. 255–257, 2011.
- [188] 平田研二, 藤田政之. 外部入力を有する線形離散時間システムに対する拘束条件の解析. 電気学会論文誌 C, Vol. 118, No. 3, pp. 384–390, 1998.
- [189] 山本江, 中村仁彦. 最大 CPI 集合に基づくスイッチングフィードバック制御によるヒューマノイドロボットの安定化. 日本ロボット学会誌, Vol. 29, No. 4, pp. 384–394, 2011.
- [190] 熊谷伊織. 積算状態推定に基づくヒューマノイドロボットの継続的タスク実行システムの構成法. PhD thesis, 東京大学大学院情報理工学系研究科知能機械情報学専攻, 2017.
- [191] 株式会社オオトモ. Ninebot One S2. <http://www.ninebot.jp/ones2/>.
- [192] Segway Inc. Segway Personal Transporter (PT). <http://www.segway.com/>.

- [193] 中川智皓, 森田悠介, 新谷篤彦, 伊藤智博. 立ち乗り式四輪型パーソナルモビリティ・ビークルにおける人間の立位姿勢での重心移動. 日本機械学会論文集, Vol. 82, No. 838, 2016.
- [194] 西内裕晶, 塩見康博. 乗車経験に着目したセグウェイの走行挙動特性に関する基礎的研究. 土木学会論文集 D3 (土木計画学), Vol. 68, No. 5, pp. 917–927, 2012.
- [195] 大槻有輝, 松野隆幸, 矢納陽, 見浪護. スライド式重心制御機構を有する倒立振子型車両の開発-小型車両の設計製作および姿勢制御実験-. 日本機械学会ロボティクス・メカトロニクス講演会講演論文集 (2013), pp. 2A2–R01, 2013.
- [196] 万波央, 風間恵介, 孕石泰丈, 毛利宏. 平行二輪型自走式倒立振子におけるトルクと車輪角速度の入力指令値比較. 日本機械学会論文集, Vol. 80, No. 814, 2014.
- [197] H. Inoue, F. Kanehiro, K. Fujiwara, S. Kajita, K. Yokoi, K. Kaneko, H. Hirukawa, Y. Nakamura, K. Yamene, and T. Isozumi. Platform for Humanoid Research and Development. In *Proceedings 2nd IARP/IEEE-RAS Joint Workshop on Technical Challenge for Dependable Robots in Human Environments*, pp. 131–141, 2002.
- [198] 井上博允, 比留川博久. 人間協調・共存型ロボットシステム. 日本ロボット学会誌, Vol. 18, No. 8, pp. 1089–1092, 2000.
- [199] 井上博允, 比留川博久. 人間協調・共存型ロボットシステム研究開発プロジェクト. 日本ロボット学会誌, Vol. 19, No. 1, pp. 2–7, 2001.
- [200] 野沢峻一. 操作情報推定に基づく接触力制御によるヒューマノイド全身適応環境物体操作の実現. PhD thesis, 東京大学大学院情報理工学系研究科知能機械情報学専攻, 2013.
- [201] T. Matsui. Multithread object-oriented language euslisp for parallel and asynchronous programming in robotics. In *Workshop on Concurrent Object-based Systems, IEEE 6th Symposium on Parallel and Distributed Processing*, 1994.
- [202] 松井俊浩. 幾何モデリング機能を備えたマルチスレッド並列オブジェクト指向言語 EusLisp. 日本ロボット学会誌, Vol. 14, No. 5, pp. 650–654, 1996.
- [203] 浅野悠紀. 人体模倣筋骨格ヒューマノイドにおける脚構造の構成法と動作実現. PhD thesis, 東京大学大学院情報理工学系研究科知能機械情報学専攻, 2016.
- [204] 上月豊隆. 筋骨格ヒューマノイドの関節駆動構造における機能統合構成法に関する研究. PhD thesis, 東京大学大学院情報理工学系研究科知能機械情報学専攻, 2016.
- [205] Y. Asano, K. Okada, and M. Inaba. Design principles of a human mimetic humanoid: Humanoid platform to study human intelligence and internal body system. *Science Robotics*, Vol. 2, No. 13, p. eaq0899, 2017.
- [206] M. Osada, N. Ito, Y. Nakanishi, and M. Inaba. Realization of flexible motion by musculoskeletal humanoid x201C;Kojiro x201D; with add-on nonlinear spring units. In *Proceedings of the 2010 IEEE-RAS International Conference on Humanoid Robots*, pp. 174–179, 2010.
- [207] Y. Nakanishi, T. Izawa, M. Osada, N. Ito, S. Ohta, J. Urata, and M. Inaba. Development of musculoskeletal humanoid kenzoh with mechanical compliance changeable tendons by nonlinear spring unit. In *2011 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 2384–2389, 2011.

以上

1p～ 326p 完

博士論文

平成 30 年 6 月 8 日提出

東京大学大学院 情報理工学系研究科
知能機械情報学専攻 博士課程
木村 航平