

人間機械系の相互作用特性の理解に基づく
行動学習支援に関する研究

廣川 暢一

目次

第1章 序論	1
1.1 人間機械系の相互作用に基づく行動学習支援	1
1.2 関連研究	2
1.2.1 人による機械系の行動学習支援	2
1.2.2 機械系による人の行動学習支援	3
1.3 目的	4
1.3.1 人間機械系の相互作用特性	4
1.3.2 相互作用の随伴性と一貫性	5
1.4 本論文の構成	6
第2章 物理的相互作用に基づく行動学習支援	9
2.1 事例研究の位置づけと概要	9
2.2 従来の運転支援システムの問題点	9
2.3 提案手法	10
2.3.1 力覚呈示によるステアリング操作支援	10
2.3.2 運転技量学習モデルと支援システムの分類	11
2.3.3 関連研究	13
2.4 駐車運転行動計測実験	14
2.4.1 実験概要	14
2.4.2 駐車成績の定量化	14
2.4.3 解析結果	16
2.4.4 考察	20

2.5	スケールカーシミュレータ	22
2.5.1	システム構成	22
2.5.2	経路計画手法	23
2.5.3	力覚呈示算出手法	26
2.6	支援効果検証実験	29
2.6.1	実験手順	29
2.6.2	駐車成績の解析結果	30
2.6.3	切返し位置と駐車成否の関係	31
2.6.4	切返し位置の分散に関する解析結果	32
2.6.5	考察	33
2.7	第3章まとめ	36
第3章	情報的相互作用に基づく行動学習支援	37
3.1	事例研究の位置づけと概要	37
3.2	背景	37
3.2.1	一般的な行動学習手法の問題点	37
3.2.2	関連研究	38
3.2.3	目的	39
3.3	教示者によるロボットの行動学習支援	40
3.3.1	主観評価の導入	40
3.3.2	人の教示の随伴性・一貫性に基づく教示意图推定	42
3.3.3	行動学習支援のアプローチ	43
3.4	実験環境	44
3.4.1	実験課題	44
3.4.2	教示入力インタフェース	45
3.5	評価関数と学習効率	46
3.5.1	評価関数の分類	46
3.5.2	検証実験	48

3.6	離散系における行動学習支援	49
3.6.1	離散系の強化学習への実装	49
3.6.2	検証実験	50
3.7	連続系における行動学習支援	52
3.7.1	連続系の強化学習への実装	52
3.7.2	検証実験	54
3.8	第4章まとめ	58
第4章	社会的相互作用に基づく行動学習支援	59
4.1	事例研究の位置づけと概要	59
4.2	背景	59
4.2.1	自閉症スペクトラム障害	59
4.2.2	療育(支援)ロボットの意義	60
4.3	関連研究と問題点	60
4.3.1	従来の療育手法	60
4.3.2	動物介在活動	61
4.3.3	療育(支援)ロボットに関する先行研究	62
4.4	提案手法	63
4.4.1	個人適応的ロボット介在活動	63
4.4.2	ロボットの実時間遠隔操作手法	64
4.4.3	快情動の推定手法	71
4.5	評価実験	72
4.5.1	実験手順	72
4.5.2	実験結果	73
4.6	第4章まとめ	77
第5章	結論	79

図目次

2.1	Haptic guidance for steering operation assistance.	10
2.2	運転技量学習モデルと支援システムの分類	12
2.3	Overview of an experimental measurement environment. The subjects were required to perform reverse parking from two initial positions to the target parking position.	15
2.4	Measurement of human driving behavior.	16
2.5	Results of analyzing the correlation between the parking score and the driving experience of the driver.	18
2.6	Correlation between parking score and total operation.	19
2.7	Correlation between parking score and operation with velocity change.	20
2.8	System overview of scaled car simulator.	22
2.9	Scaled car simulator. (a) Bird's eye view, (b) Front view, (c) Front view from the driving seat, and (d) Rear view from the driving seat.	23
2.10	Path-planning algorithm based on Bezier curve.	24
2.11	A flow chart of haptic instruction calculation.	27
2.12	Comparison of averaged parking scores of the groups.	31
2.13	Examples of parking trajectories. (a) subject in group A, (b) subject in group B.	33
2.14	Absolute variance of the direction change point.	34
3.1	Behavior learning assistance based on human subjective cues.	40
3.2	Causality detection and Error detection.	40
3.3	Behavior learning assistance based on human subjective cues.	41
3.4	Time-consistency based human intention estimation.	42
3.5	Approach of behavior learning assistance for reinforcement learning agent.	44
3.6	6 DOF robot arm.	45

3.7 Simulation model	46
3.8 The input interface for teaching real robots.	47
3.9 Reward functions for the validation experiment.	48
3.10 Results of the validation experiment.	49
3.11 Results of the comparison experiment in a discrete environment. . .	51
3.12 The transition of the reward function.	52
3.13 Results of the simulation experiment.	55
3.14 The transition of the reward function in the simulation experiment.	56
3.15 Results of the experiment with a real robot.	57
4.1 Animal-Assisted Activities	61
4.2 Proposed Robot-Assisted Activity.	63
4.3 A doll-type robot control interface.	64
4.4 The result of the comparison experiment between the proposed doll- type interface and conventional GUI based interface.	65
4.5 Demonstration of the balancing algorithm.	68
4.6 Demonstration of motion pattern recognition.	70
4.7 A wearable device for facial expression detection.	71
4.8 Experiment field.	73
4.9 Experiment overview (Normal control participant)	74
4.10 The percentage of time the subject performed positive social behav- iors in animal-/robot-assisted activity.	75
4.11 The percentage of time the subject performed negative social behav- iors in animal-/robot-assisted activity.	76
4.12 Time percentage of Face-to-Face behaviors related to positive social behaviors.	77

表 目 次

1.1	Behavior learning characteristics of human and machine	2
2.1	Criteria for scoring the parking.	16
2.2	κ coefficients among human coders (Position/Orientation)	17
2.3	Results of correlation analysis	17
2.4	Conditions of experiment	30
2.5	Relationship between the direction-change point (DCP) and the rate of successful parking	32
4.1	Leg states and the recognized motion	69

第1章 序論

1.1 人間機械系の相互作用に基づく行動学習支援

ロボット技術の発展により、近年高度な自律性と対人親和性の高いインタフェースを備えた機械システムが一般社会へと急速に普及しつつある。それに伴い、人と機械の相互作用 (Human-Machine Interaction) に関する研究も従来の限られた環境での単純な作業分担や作業代替を目的としたものから [1], 人や機械系の自律的な行動学習メカニズムを明示的に取り入れ、人間機械系の相互作用を通じて行動の獲得や洗練、拡張を目指す試みが増えてきている。

ここで、人の行動学習は一般に環境の認知に基づく行動の選択とその結果の評価を繰り返すことで行う。このような人の行動学習メカニズムを抽象化し、計算機で再現する試みが強化学習に代表される機械系の行動学習理論であり [2], ロボットに代表される大規模で複雑な計算に基づく学習メカニズムを実装した機械システムが日常生活の様々な場面に浸透しつつあることから、機械系の自律的な行動学習手法の重要性が高まっている。しかしながら、環境認知の不確実性や次元数の増加に伴う探索空間の増大やフレーム問題、Heuristic な評価関数設計など、既存の行動学習手法を実問題へ応用するために解決すべき多くの課題が残っている。一方、人は高次元の複雑な状況認知や抽象的・主観的な価値に基づく判断など、機械には非常に困難な処理を素早く正確に行うことができるが、高速で精密な動作や計算速度では機械には及ばず、認知的・身体的な入出力特性による制約も持つ。

このように、人や機械システムの行動学習特性にはそれぞれ長所と短所を持つことから (表 1.1), 人がシステムの制御系内に物理的・情動的に介入することで人の持つ能力や知見を機械系の行動学習に活用することや、機械系の支援によって

表 1.1: Behavior learning characteristics of human and machine

	Human	Machine
Advantages	<ul style="list-style-type: none">• Complex and abstract cognitive skill• High inference ability	<ul style="list-style-type: none">• Fast and precise movement• High calculation ability
Disadvantages	<ul style="list-style-type: none">• Cognitive and physical constraints• Weak calculation ability	<ul style="list-style-type: none">• Disability of complex and abstract cognition• No-free-lunch theory

人の行動学習過程を促進するといったアプローチが考えられ、本研究ではこれを人間機械系の相互作用に基づく「行動学習支援」と定義する。具体的には、複雑な環境下での行動学習に必要な認知的処理を人が補うことで機械系の行動学習を支援することや、一方で人が適切な行動を直観的に判断しづらい問題や、障害などによって身体的な動作に制約のある人のための、システムによる人の運動の物理的な支援手法などがこれまでに提案されている。

1.2 関連研究

1.2.1 人による機械系の行動学習支援

人が機械系の行動学習を支援するためのアプローチは二通りに大別できる。一つは、ロボットが獲得すべき行動を人が例示することで学習を支援するアプローチである [3] [4] [5]。これによってロボットによる行動空間の探索時間を大幅に削減することが出来るとともに、人によるロボットの行動プログラミングに関する負担の軽減にもつながる [6]。Atkeson らは例示によるロボットの学習アルゴリズムを考案し、倒立振子の振り上げと安定化タスクによってその有効性を示した [7] [8]。Inamura らは例示によって与えられた多次元の関節空間上の時系列パターンを隠れマルコフモデルによってシンボル化し、ロボットの全身行動の認識と生成を同

時に行うことの出来る手法を提案した [9] [10]. また, 人が直接ロボットの身体部位を動かすことで動作をプログラミングするなど, ロボットの動作設計に関する専門知識を必要としない手法や [11], 成功例ではなく失敗例を示すことで学習を支援する手法なども提案されている [12].

一方, 行動学習における評価のプロセスを人が代行することで学習を効率化させる手法も多く提案されている [13]. Nakatani らは人がロボットに対して主観的な評価をオンラインで与える Coaching 法を提案し, 二足ロボットの歩行やバランス制御のためのパラメータ探索を事前の Fine-Tuning 無しに行えることを示した [14]. Reliy らは Nakatani らの結果を基により具体的な複数の主観評価を導入した Coaching 法を提案し, 人型ロボットの動作パラメータ調整を直観的に行える手法を開発した [15]. また筆者らも, 人の主観評価に基づき強化学習エージェントが持つ評価関数を動的に更新することでロボットの行動学習を支援する手法を考案し, 実験によってその有効性を示してきた [16] [17]. その他にも, 進化計算手法に主観評価を取り入れることで, 定量化の困難な人の主観的な好みをモデル化する試みや [18], 主観評価の入力方法として装着型インタフェースを用いて検出した人の表情を用いる手法なども提案されている [19].

1.2.2 機械系による人の行動学習支援

機械系による人の行動学習支援としては, 機械システムの得意とする正確で繰り返しの容易な動作制御を用いて事故や病気で身体機能に障害を抱えた人の運動機能の再獲得を支援するシステムが多く開発されており, リハビリテーションの現場で既に実用的に用いられているものも少なくない. Aisen らは, 上肢のリハビリテーションのための MIT-Manus を開発し, 精密なフィードバック制御に基づき使用者の自発的な動作の補助・誘導することで, 効果的なリハビリテーションを行うことが出来ることを示した [20]. これは, 従来の人間の理学療法士によるリハビリテーションでは実現が困難な運動であり, 同じ訓練を長時間実施出来るなど, 機械システムによるリハビリテーション支援の可能性を示唆する重要な成果

である。また、下肢の運動機能を支援するシステムとして外骨格型の支援機器などがある [21] [22]。Sankai らは Robot suit HAL を開発し、下半身の運動機能に障がいを持つ人のための自立歩行やリハビリテーションの支援を行えることを示した [23] [24]。リハビリテーションの観点から見た HAL の最大の特徴は、筋電位信号に基づいて使用者の運動意図を推定し、実際の四肢の運動と同じタイミングでの支援機器の制御を可能にしている点である。これによって、自発的な運動意図に同期した体性感覚のフィードバックにより脳の神経可塑性を促進することで、効果的な運動療法を自動でかつ継続的に実施できる [25]。

また近年では、発達障害を持つ児童との社会的なインタラクションを通じて、社会性の向上やソーシャルスキルの学習を支援する療育ロボットが注目されている [26] [27]。これらは、主に自閉症スペクトラム障害 (Autism Spectrum Disorders, ASD) を持つ児童 (ASD 児) 対象としており、ロボットを用いることで多くのメリットが得られることから、様々な療育 (支援) ロボットがこれまでに提案されている。Robins らは、人間の子供と似通った外観を持ち様々なインタラクションを行えるロボットを開発し、遊びを通じた社会性の向上や複数の ASD 児間のインタラクションを促すことを目的とした臨床実験を行っている。また Pioggia らは表情を用いる非言語的なコミュニケーションスキルの学習支援を目的として、様々な表情を表出できる人に酷似したロボットを開発している [28]。

1.3 目的

1.3.1 人間機械系の相互作用特性

上記のように、人間機械系の相互作用によって人または機械系の行動学習の支援を目的とした取り組みは多く実施されているが、行動学習支援のためのシステム設計やインタラクションのデザインに関する統一的な方法論は確立されておらず、現状では実装する人間機械系のシステムに応じて試行錯誤的な設計・開発が行われている。そこで本研究では、環境やインタラクション形態に依存しない人

と機械システムの相互作用に関する一般的な特性の理解に基づき、様々な人間機械系のシステムにおける行動学習支援の実現を目的として研究を行う。ここでは、人と機械系のインタラクションの様相という観点から、先行研究を1) 物理的相互作用に基づく行動学習支援、2) 情動的相互作用に基づく行動学習支援、及び3) 社会的相互作用に基づく行動学習支援に分類し、関連する実問題についてそれぞれ行動学習支援システムの提案と開発を行う。これによって、人間機械系の相互作用に基づく行動学習支援という観点から共通する性質やシステムの要件を明らかにすることで、人間機械系の行動学習支援のための統一的な方法論に関する知見を得るとともに、得られた知見をフィードバックすることで既存の機械学習手法の改善や、行動学習に関する人の特性についての更なる理解に貢献することも期待出来る。

1.3.2 相互作用の随伴性と一貫性

行動学習支援における重要な人間機械系の相互作用特性として、ここでは相互作用の随伴性 (Contingency) と一貫性 (Consistency) を考える。相互作用の随伴性とは、行動学習過程への介入とそれに対する応答が実時間で非同期的に行われる特性を意味し、任意の時点で必要な介入を行うことや、介入による学習への影響を時間的な近接性に基づく因果関係の推論を用いて評価するために必要である。また、一貫性とは相互作用の文脈に時間的な整合性が保たれている特性のことであり、やり取りされる情報の信頼性を評価することにつながると考えられる。これらは、人や動物の行動学習メカニズムを主体的な行動に随伴する報酬または罰を用いて説明した行動分析学 [29] や、乳幼児の発達過程を愛着という観点から分析したアタッチメント理論 [30] から着想を得ており、報酬/罰の随伴性と学習効率の関係や一貫性の無いフィードバックによって学習効果が減衰することなども知られている。これは人の行動学習に関する一般的な特性であるといえ、また人の行動学習メカニズムを模倣して考案された強化学習 [2] などの機械系の行動学習手法にも同様に当てはまることが予想される。

1.4 本論文の構成

本論文の構成は次の通りである。

第2章では、人と機械の物理的な相互作用に基づく行動学習支援の事例研究として取り組んだ、後ろ向き駐車運転の運転技量を育成する運転支援システムについて述べる。ここでは、後ろ向き駐車に関する人の運転特性と駐車技量との関係を明らかにするため実車を用いた実験を行い、それによって得られた知見に基づき力覚呈示によってステアリング操作を実時間で支援するシステムを開発するとともに、このシステムを用いることで人の運転技量の習熟を有意に促進出来ることを被験者実験によって示す。

第3章では、情動的相互作用に基づく行動学習支援の事例研究として、強化学習を実装したロボットに人が教示者として直感的で簡便な教示を実時間で与えることで行動学習を支援するためのアルゴリズムを考案する。その際、人の認知判断に関する誤りや時間遅れといった認知的特性の理解に基づく、頑健な人の教示意図推定と報酬関数の更新手法を考案する。この手法により、既存の強化学習による学習効率を改善するとともに、設計者による報酬関数設計の負担を削減出来ることを実験によって明らかにする。

第4章では、社会的相互作用に基づく行動学習支援の事例研究として、自閉症スペクトラム障害をもつ児童のためのロボット介在活動について述べる。これは、自閉症児の社会的行動と情動の関係に着目し、快情動を誘発することで他者に対する自発的なコミュニケーションを促すことで社会的行動の学習を支援する試みである。ここでは、自閉性傾向や現有スキルの様々に異なる個々の児童に合わせたロボット介在活動を実施するための環境として、療育者のための直観的なロボットの行動制御手法を開発する。同時に、表情筋活動から笑顔を検出する装着型表情計測デバイスや社会的インタラクションを定量的な計測による、個人適応的なロボット介在活動のフレームワークを構築するとともに、提案する行動学習支援の可能性を実験によって検証する。

最後に、以上の事例研究を通じて得られた人間機械系における行動学習支援の

ための相互作用特性についてまとめと考察を行うとともに，人にとって直観的で察しのよい機械システムを実現するための人間機械系の相互学習に基づく個人適応制御に関する今後の展望を第5章で述べる．

第2章 物理的相互作用に基づく行動学習支援

2.1 事例研究の位置づけと概要

本章では、物理的相互作用に基づく行動学習支援の事例研究として駐車運転技量育成システムについて述べる。これは、自動車の後ろ向き駐車に関する人の運転技量を、ステアリングへの力覚呈示によって向上させるためのシステムであり、機械システムによる介入が人の行動学習過程に与える影響の調査を目的として行った。

2.2 従来の運転支援システムの問題点

自動車の運転は日常生活で最も身近な人間機械協調系の一つであり [31]、計算器の小型化・高性能化、センサ・制御技術の発展を背景として運転者への実時間情報呈示や部分的な自動化による運転支援技術が数多く開発されている。具体的には、運転者の状態を計測することで危険を予測し注意を促す、または車両に搭載したセンサを用いた環境認知に基づく制御により危険を自動的に回避するといったシステムがこれまでに提案されてきた [32] [33] [34]。しかしながら、運転者への視聴覚情報提示や車両の自動制御に基づく従来の運転支援システムでは、過度な依存による運転への注意力低下や [35]、自動制御による運転者が予期しない車両の動作に対する驚き・恐怖 (オートメーション・サプライズ)、システム作動中に人的物的な被害が発生した場合の責任の所在といった、悪影響や問題点も指摘される [36]。一方、人と機械系の協調作業を必要としない完全自動運転が可能な車の実用化には、あらゆる場面での安全な走行を実現するための技術的な困難だけ

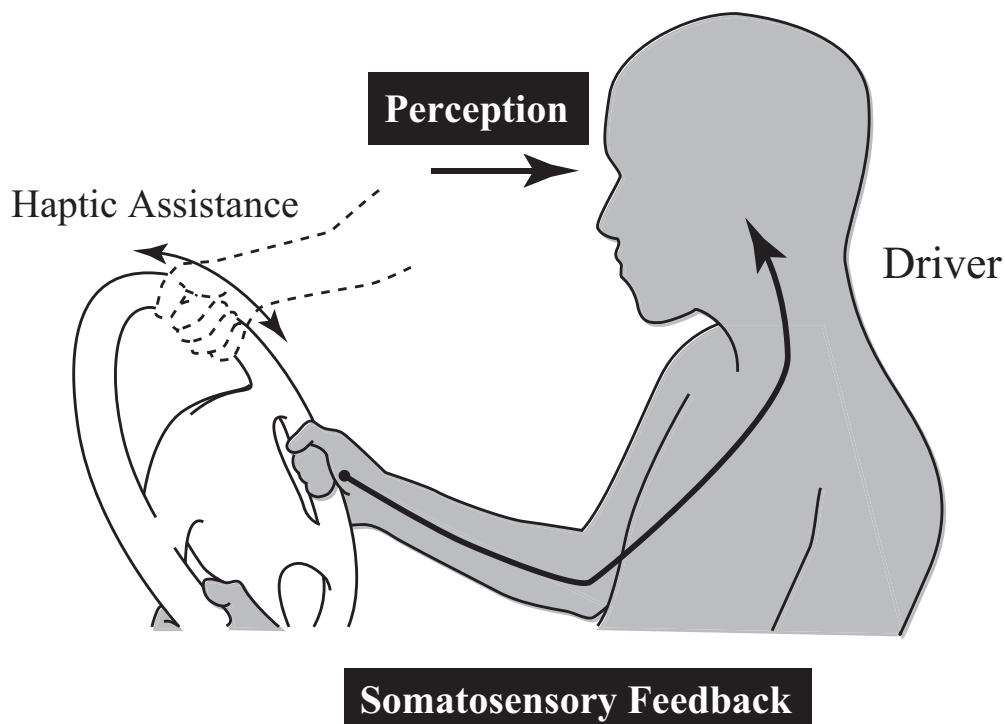


図 2.1: Haptic guidance for steering operation assistance.

でなく、必要な環境インフラの整備・普及といった法制度的な観点からも当分時間がかかることが予想される。また、安全性の面からもステアリングやペダルといった旧来のインタフェースが撤廃されることは考えにくいことから、自動運転システムに頼らない運転技術の習得は依然として重要である。

2.3 提案手法

2.3.1 力覚呈示によるステアリング操作支援

そこで本研究では、車を運転する人の運転技量に着目し、システムによる人の運転行動への適切な介入によって運転技量の育成を目的とする新しい運転行動支援システムを提案する。ここでは、様々な運転シーンの中でも比較的難易度が高く、かつ日常的に必要となる場面の多い後ろ向き駐車を対象とし、運転技量育成

のための支援システムを開発する。具体的には、図 2.1 に示すようにステアリングを介した力覚呈示により適切な操作を直接的に運転者に教示する手法を提案する。これによりステアリングの操作量やタイミングを運転者が直感的に認識することが出来るとともに、呈示される力に合わせて運転者自身が意識的に操作を行うことで、目標とする運転行動の効率的な学習が出来ると期待される。また、腕部動作であるステアリング操作と同じ感覚様相へ直接的に教示するため、従来の視聴感覚への情報提示に比べて認知的負荷が少ない支援が可能となるとともに [37]、視覚・聴覚情報を用いる他の支援手法との併用も用となる [38] [39]。さらに、目標角度からの誤差に応じた反力を呈示する Soft-Guidance [40] 法を採用することで高い学習効果を得られるとともに、人の主体的なステアリング操作を阻害しない程度の支援強度にすることでシステムの予測しない緊急時の安全性も確保できる。

提案手法と同様に運転者の技量向上によって乗り心地の良い運転や燃費の低減を目的としたシステムとして、マツダ自動車が開発したインテリジェントドライブマスタがある [41]。これは、運転者の操作情報や車両情報から走行中の車の挙動や乗車員の体の動きをモデル化することで、現在の運転に対し「丁寧」「荒い」「しなやか」といった 3 段階の評価を実時間で視覚的に運転者に呈示するシステムである。しかしながら、このシステムによって運転者に呈示されるフィードバックは簡便な視覚情報のみであるため、様々な状況に応じた同時かつ複合的な操作から成る運転操作を効率的に学習するには不十分である。

2.3.2 運転技量学習モデルと支援システムの分類

自動車の運転行動は図 2.2 に示すように、「認知」→「判断」→「操作」という一連の要素から成り、運転者による操作入力（ステアリングやアクセル・ブレーキ）を介し、最終的に車両の「移動」として運転者に知覚される。ここでは、従来提案されている様々な運転支援システムにおける支援形態を分類する観点として、それぞれの手法が対象とする運転行動の要素に着目し、それぞれ「認知支援」「操作支援」「移動支援」とする。

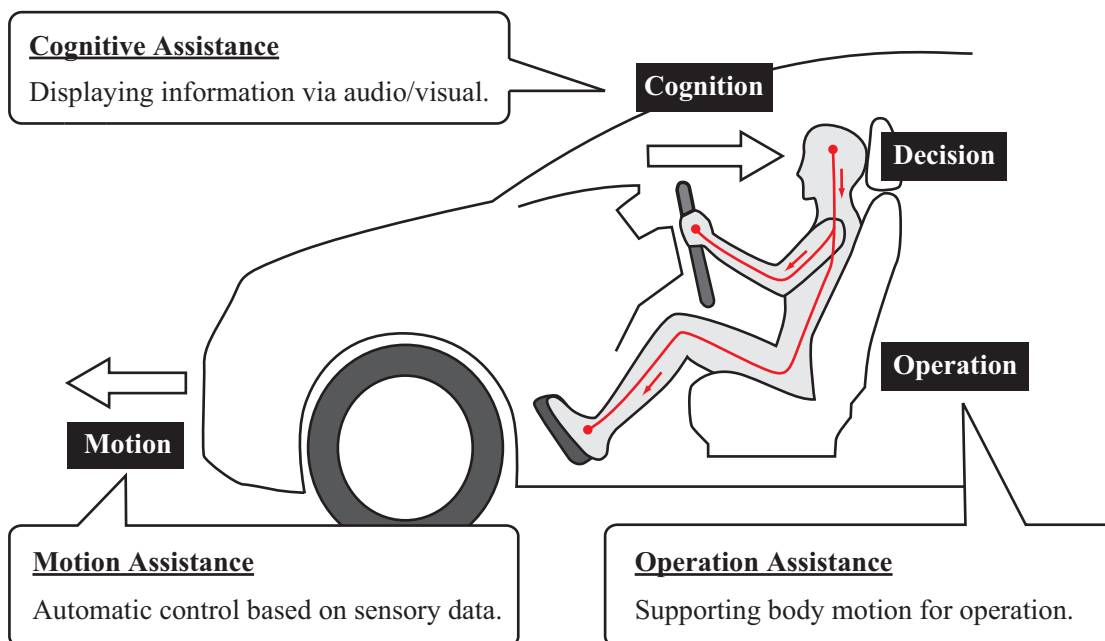


図 2.2: 運転技量学習モデルと支援システムの分類

認知支援とは、車両の状態や周辺の道路状況等、運転者単独では計測が困難もしくは不可能な情報を視聴覚情報として呈示することで、運転に係る認知行動の支援を目的とした手法を指す。加藤らは車両に搭載したセンサや車車間通信によって得られた周辺情報に基づき、操作の誘導や操作量を Head-Up Display(HUD) に表示するシステムを提案している [42]。また、現在の舵角や速度から求められる自車両の予測位置を HUD に提示することで、運転者の認知支援を行うシステムも開発されている [43]。その他にも、後退中に車体の後方を表示するバックモニター、車体周囲に取り付けたカメラ画像を合成して上空からの視点を再構成する鳥瞰モニター [44] など一般的になりつつある。しかしながら、一般的な認知支援システムでは呈示された情報に基づく操作の判断は運転者自身に委ねられており、そのため運転経験の未熟な運転者では認知支援技術によって拡張/付加された情報に対し常に適切な判断を行えるとは限らない。また視覚的な認知支援の場合、呈示画面に集中することで本来の運転への注意が疎かになり、事故の可能性が増すといった危険性も指摘されている [45]。

一方、移動支援とは運転者による操作の結果や車両制御システムの出力としての「移動」を支援するための技術を指し、車体に搭載したセンサ情報に基づく車両の自動制御技術などがこれに相当する。本研究に関連する具体例として、駐車場への自動駐車を実現するインテリジェント・パーキングアシスト [46] やスマートパーキングアシストシステム [47] などがあり、同様のシステムは他にも複数開発・販売されている。上記のシステムは、いずれも駐車に関する制御は基本的にシステムが全て自動で行うため運転者が関与する余地は限られており、動作中のシステムに対して運転者が介入した場合には即座に解除されるように設計されている。したがって、これらのシステムの利用による運転者自身の駐車技量の向上は見込めず、逆にシステムへの過度な依存による運転への注意力低下などの悪影響も懸念される [48]。

また操作支援とは、運転に係る人の身体動作に対する支援の枠組みであり、適切な操作方法を運転者に教示する手法や、運転者の主体的な操作の補助を行う手法を指す。本研究における運転技量の習熟とは、運転者が現在の状況を認知し、それに適した操作を導く認知・判断・行動から成る一連の過程の確立・洗練・簡略化を指し、一般に運転技量は図 2.2 に示すフィードバック系での反復練習によって向上すると仮定する。そのため、運転技量の習熟速度を支援によって増進させるには認知支援や移動支援のみでは不十分であり、提案手法のような運転者による主体的な操作を促す「操作支援」が不可欠であるといえる。

2.3.3 関連研究

力覚呈示を用いた運転支援システムはこれまでも複数提案されており、実証実験を通じてその有効性も報告されている [49]。Griffiths らは、ステアリングへの力覚呈示による運転支援によって運転者の軌道追従や障害物回避の成績が向上するとともに、視覚的な認知に関する負担を軽減出来ることをシミュレーション環境を用いて示した [37]。また Saleh らは、人の運転動作の不確実性を考慮した運転モデルを考案し、それに基づくレーンキーピングのための Shared-control 手法を

提案した [50][51]. その他にも, 力覚呈示に対する人の反射動作を利用して運転中の車線離脱を警告するシステムや [52], カーブ走行時の安定性を向上させるためのアシストシステム等も開発されている [53]. ステアリング操作以外の運転操作に対しても, ガスペダルを用いた力覚呈示による操作支援手法が提案されている [54][39]. しかしながら, これらの先行研究ではシステムの使用における安全性や安定性の評価はなされているものの, システムを使用することによる運転者本来の運転技量への影響を評価している取り組みは少ない [55]. そのため, 力覚呈示による操作支援システムによる人の運転行動の学習過程への影響を調査することで, 人の適応メカニズムを考慮に入れた支援モデルの構築などが期待される.

2.4 駐車運転行動計測実験

2.4.1 実験概要

はじめに, 適切な支援方法の検討のため, 後ろ向き駐車時の人の運転特性の計測を目的とした実験を, 図 2.3 に示すような環境で市販のワゴンタイプの車に計測機器を搭載した実験車両を用いて行う. ここでは, 図中斜線で示した駐車スペースに対し左右二ヶ所に開始位置を設定し, 駐車スペースの向きに対して直交する位置で静止した状態からの駐車行動を計測した. また, 駐車時の切り返し回数は 1 回までとし, 一人の被験者に対して左右の開始位置から交互に 5 回ずつ, 計 10 回計測する. 被験者は普通自動車の運転免許を持つ男女 10 人 (男性 9 人, 女性 1 人, 運転歴 4~30 年) とし, 事前に実験内容についての同意を得た上で行う. 本実験で計測するパラメータとその計測方法は表 1 の通りである. また, 駐車成績の評価のため実験環境を俯瞰する位置からビデオカメラによる撮影も同時に行った.

2.4.2 駐車成績の定量化

計測結果の解析にあたり, 始めに各被験者の駐車成績を, 駐車完了時の車体の位置・向きおよび駐車に掛かった時間に基づいて定量化を行う. 駐車的位置・向き

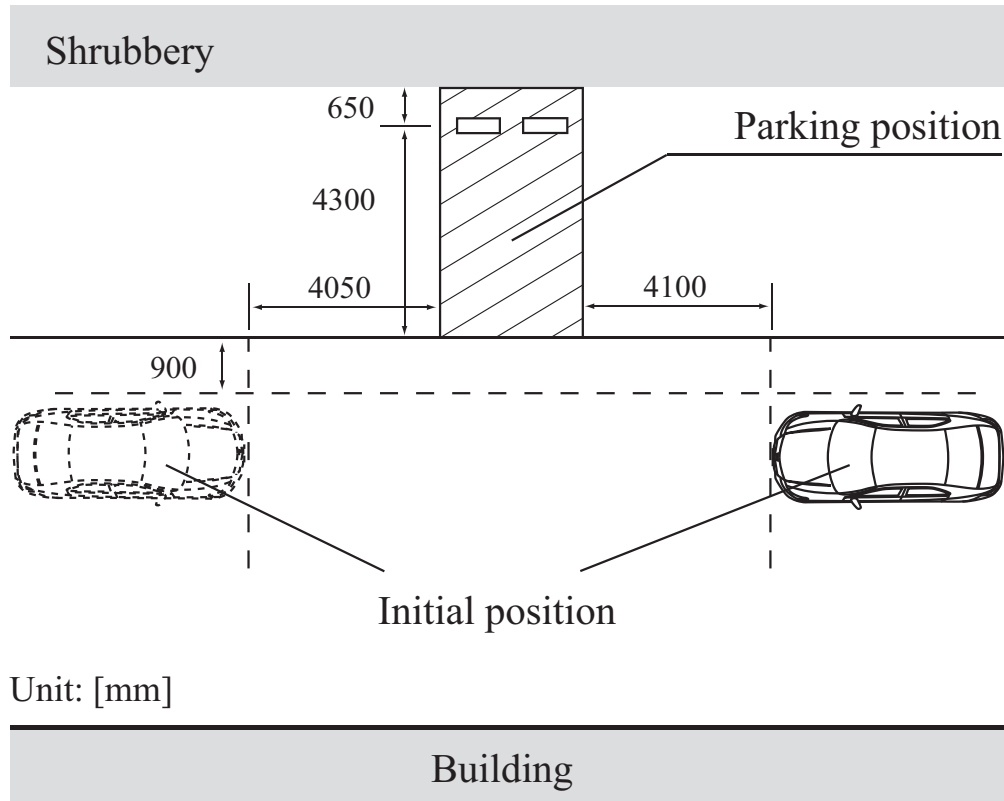


図2.3: Overview of an experimental measurement environment. The subjects were required to perform reverse parking from two initial positions to the target parking position.

は撮影したビデオ映像から目視による4段階のリッカート尺度で評価する。評価基準は駐車的位置・向きに関して表2.1に示すように単純かつ客観性の高い指標を用い、ここでは自動車の駐車行動について熟知している三人の実験者が独立に評点化し、それらの平均を各試行の評価値とした。図2.4に、実際の駐車結果とその時の評点を例示する。

表2.2は評点者間の評価の一致度を測るために行った、評点者間の κ 係数をまとめたものである。この表から、評点者間での κ 係数は最低でも0.889と高い値となっており、このことから同一の試行に対する各評点者による評価はほぼ一致していると考えられ、駐車精度を表す信頼性の高い客観評価値として解析に用いることは妥当であると判断する。また、切返し位置から駐車完了までに要した時

表 2.1: Criteria for scoring the parking.

Score	Position	Orientation
4	Center of parking area	Straight
3	Inside lines but lopsided	Slightly tilted
2	One or more wheels on the line	Tilted
1	One or more wheels beyond the line	Very tilted

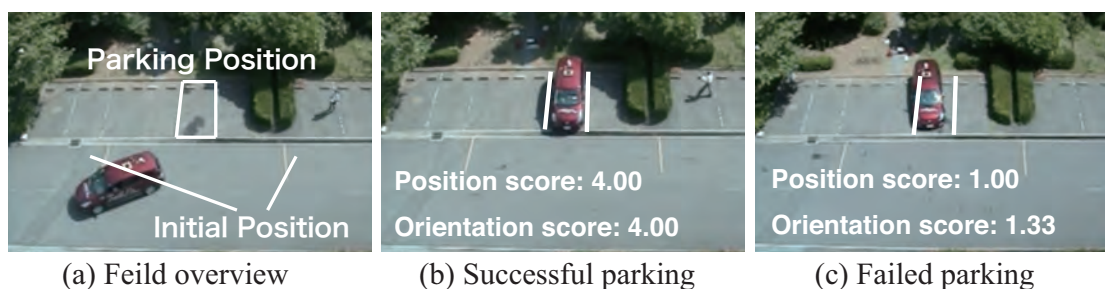


図 2.4: Measurement of human driving behavior.

間も合わせて計測し、駐車位置に関する評価と足し合わせることで、どれだけ早く正確に駐車出来たかを定量的に示す駐車成績として定義する。なお、両評点は平均3、標準偏差1となるよう正規化を行った上で足し合わせる。今回は10人の被験者の内、映像の欠損により駐車位置の評価が出来なかった1名を除いた、9名×10回分の駐車データを用いて以降の解析を行った。

2.4.3 解析結果

始めに、駐車成績と運転歴の相関解析を行った結果を図 2.5 に示す。図の横軸は運転歴を表し、縦軸は10回分の駐車成績の平均と標準偏差をプロットしたものである。この図から、運転歴と駐車成績には緩やかな正の相関が見られ、熟練運転者と考えられる運転歴の長い被験者ほど、駐車成績も高い傾向があることが分かる。ここで、運転歴が17年、20年、26年の被験者は普段から実験車両と同様のワゴンタイプを運転しているが、図 2.5 の結果からは同程度の運転歴で異なるタイプ

表 2.2: κ coefficients among human coders (Position/Orientation)

	Experimenter A	Experimenter B	Experimenter C
A	1	0.935 / 0.925	0.889 / 0.917
B	0.935 / 0.925	1	0.926 / 0.953
C	0.889 / 0.917	0.926 / 0.953	1

表 2.3: Results of correlation analysis

	# of samples	Correlation coefficient	$P(T \leq t)$
Total operation	90	-0.569	2.35×10^{-9}
Operation with velocity change	90	0.695	1.05×10^{-14}

の車に乗っている他の被験者との間に明確な駐車成績の差は見られなかったため、実験車両の特性による駐車成績への影響は少なかったと考えることができる。

次に、人の運転操作に関する定性的な知見を得るため、運転時のステアリング、アクセル及びブレーキ操作の計測データの解析を行う。ここでは、1回の試行における各操作の総量を以下のように定義する。

$$\Delta x(t) = |x(t) - x(t-1)| \quad (2.1)$$

$$M = \frac{1}{\max x} \sum_{t=2}^T \Delta x(t) \quad (2.2)$$

M は正規化した操作量を表し、 $x(t)$ は 1 試行分の計測データ、 T はデータ数を表す。また、ステアリング操作量を M_{str} 、アクセル操作量を M_{acc} 、ブレーキ操作量を M_{brk} とし、それらの合計を全操作量 M_{total} として以下のように定める。

$$M_{total} = M_{str} + M_{acc} + M_{brk} \quad (2.3)$$

これらはいずれも車体の CAN ネットワークから 100 [ms] 周期で直接取得した値を

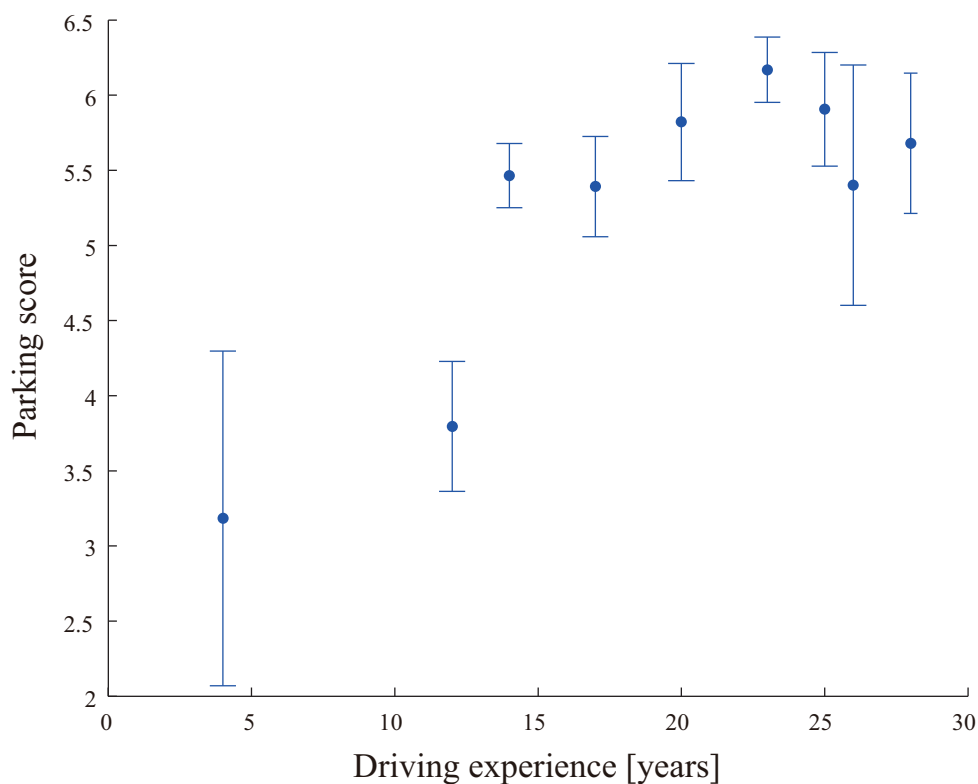


図 2.5: Results of analyzing the correlation between the parking score and the driving experience of the driver.

用いる。

図 2.6 及び表 2.3 の上段に、駐車成績と全操作量の相関解析結果を示す。この結果から、駐車成績と全操作量の間には負の相関が見られ、t 検定によって求めた相関関係の危険率も非常に低い確率を示した ($p \leq 0.001$)。

次に、駐車に要した時間の観点から駐車中の車両の速度に着目し、加速度として運転者に認知される車両の速度変化とその時のステアリング操作量について解析を行った。ここでは、車速とステアリング角が同時に変化したときのステアリング操作量 $\Delta x'_{str}$ 及び速度の変化量 $\Delta x'_v$ の和を、車速の変化を伴うステアリングの操作量と定義し、加減速時操作量 M_{vc} と呼ぶ。

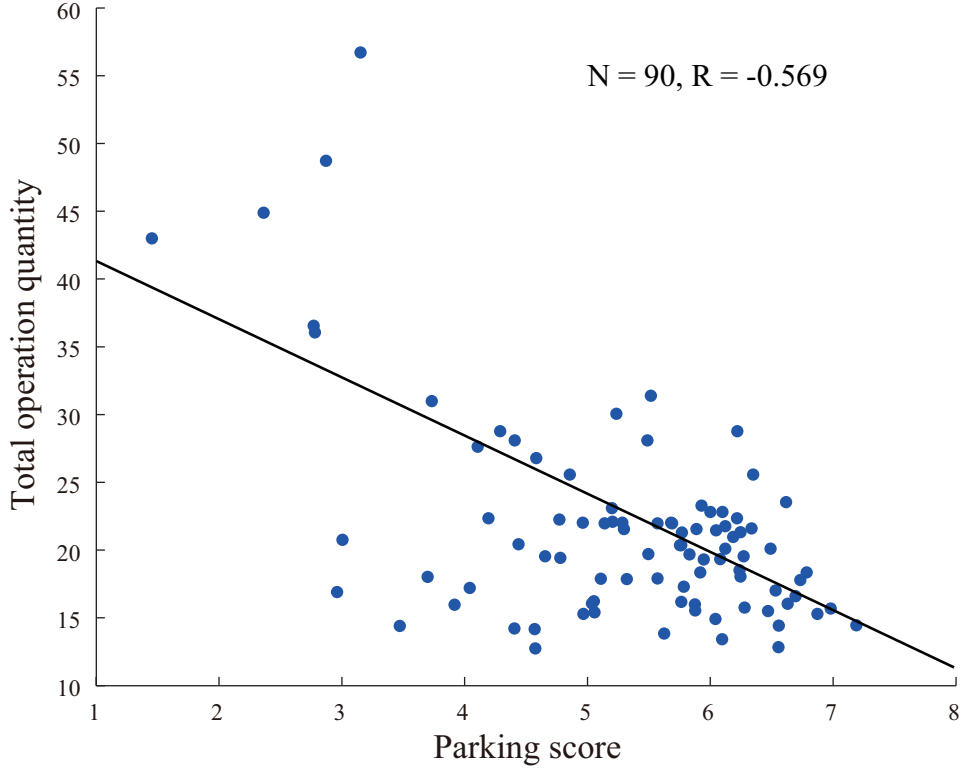


図 2.6: Correlation between parking score and total operation.

$$\Delta x'_{str}(t) = |x_{str}(t) - x_{str}(t-1)| \quad [\Delta x_v \neq 0] \quad (2.4)$$

$$\Delta x'_v(t) = |x_v(t) - x_v(t-1)| \quad [\Delta x_{str} \neq 0] \quad (2.5)$$

$$M'_{str} = \frac{1}{\max x_{str}} \sum_{t=2}^T \Delta x'_{str}(t) \quad (2.6)$$

$$M'_v = \frac{1}{\max x_v} \sum_{t=2}^T \Delta x'_v(t) \quad (2.7)$$

$$M_{vc} = M'_{str} + M'_v \quad (2.8)$$

ここで、加減速時操作量の全操作量に対する割合を加減速時操作率と定義し、駐車成績との相関解析を行った結果を図 2.7 及び表 2.3 下段に示す。この結果より、駐

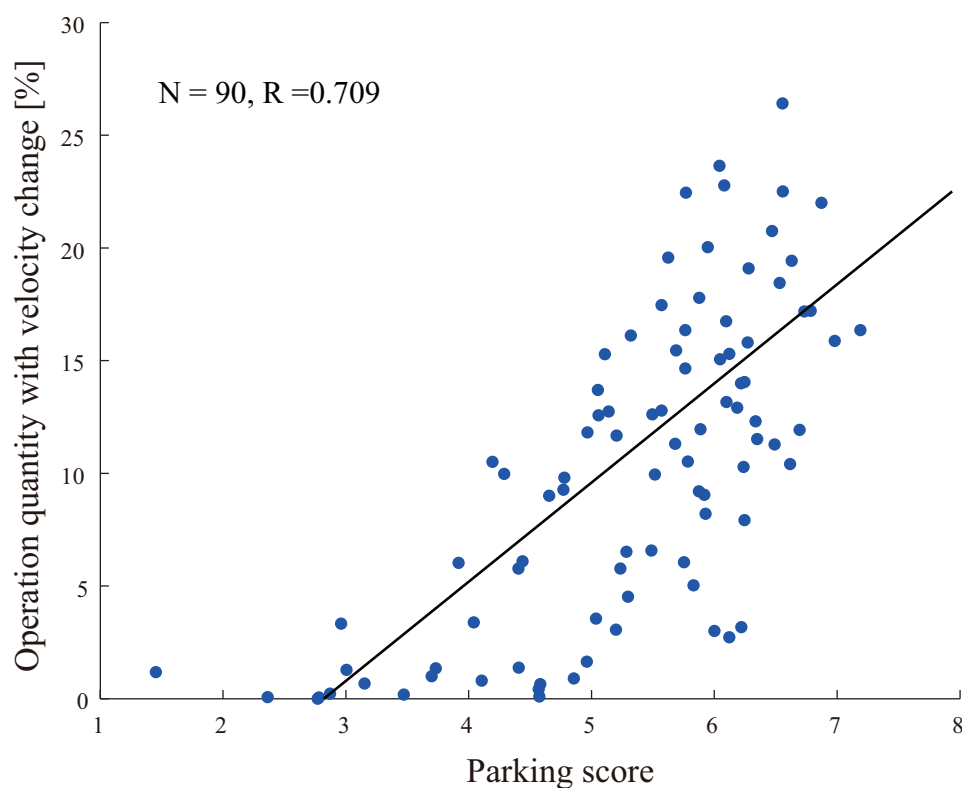


図 2.7: Correlation between parking score and operation with velocity change.

車成績と加減速時操作率の間には強い正の相関が見られ、その有意確率も t 検定において高い値を示した ($p \leq 0.001$)。

2.4.4 考察

以上の解析結果より得られた人の運転特性に関する知見として、まず運転歴と駐車成績に正の相関があることから、運転経験の蓄積によって運転技量が向上するという直感的な推測とは矛盾しない結果が、実際の被験者実験からも得られたといえる。しかしながら、この場合の運転歴は免許取得後の経過年数であり、車の取得時期や普段の運転頻度によって必ずしも運転技量を正確に表すとは言えないため、今回は駐車成績という実際に行った駐車の結果のみから算出される客観

的な指標を用いて評価を行った。

駐車成績と運転操作に関しては、まず全操作量と駐車成績に負の相関が見られることから、駐車成績の高い運転者ほど修正舵等無駄の少ない操作で駐車を成功させていると考えられ、これは一般的な運転技量に関する事前の予測と一致する。また、加減速時操作率と駐車成績についても正の相関関係が確認できたことから、駐車成績の高い試行ほど変化する車速に合わせて適切にステアリング操作を行うことで、素早く正確に駐車をを行っていることが分かる。逆に、駐車成績の低い試行では、車速が一定、もしくは停止している時にステアリングを切る、所謂「据え切り」を行っている割合が高いと考えられる。この理由を説明する一つの仮説として、定速でのステアリング操作による車両の挙動は比較的容易に予測出来るが、車速の変化を伴うステアリング操作は認知的な負荷が高いため、運転に不慣れな被験者ほど車両の速度を一定に保ちながら駐車をしているためではないかと考えられる。これによって、ペダル操作を伴う車体の速度制御時においては適切なステアリング操作が行われず、結果として駐車が失敗する可能性が高まると考えられ、熟練ドライバーの運転分析に関する先行研究でも同様の報告がなされている [56]。このことから、車速が変化する中でも滑らかに操舵を行い、かつ全体としての操作量が最小となるような駐車行動が運転技量の高い駐車と言える。そこで、得られた上記の運転特性を実現するための支援手法として、運転者の任意の速度においても逐次最適制御に基づき適切に操作教示を行う運転行動支援のための力覚呈示手法を提案する。この手法は、最適制御を用いることで駐車の成功を補償するとともに、操作量および所要時間の最小化等、本実験より得られた人の運転特性と運転技量の関係の点からも妥当な手法といえる。

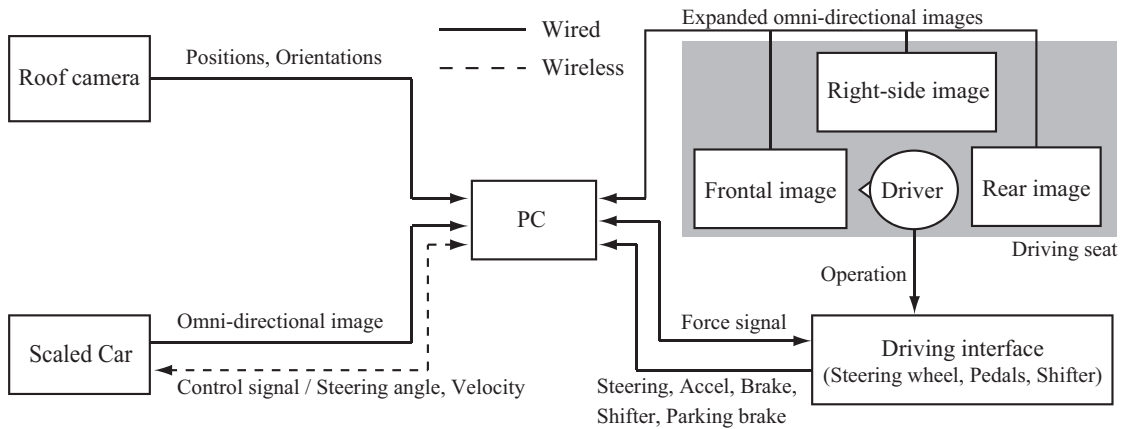


図 2.8: System overview of scaled car simulator.

2.5 スケールカーシミュレータ

2.5.1 システム構成

図 2.9 に、提案手法の実装するためのプラットフォームとして開発したスケールカーシミュレータの概要を示す。これは、実車の 1/10 スケールのモデルカーの運転席部分に小型の全方位カメラを搭載し、得られた画像を運転席周囲に配置した画面に映し出すことでモデルカーの運転席視点での運転体験を可能とするシステムである (図 2.8)。運転席にはステアリングとペダル、シフトレバーなど通常の車両と同じインタフェースを実装しており、モデルカーは運転者の操作入力に基づき無線通信で遠隔操作される。また、バックミラー・サイドミラー等の内・外装も実装しており、周囲に実環境を模したジオラマを配置することで異なる運転環境を再現することや、実画像を用いることによる臨場感のある運転シミュレーションが可能となる。車両および駐車目標位置は実験環境を俯瞰するように設置した天井カメラによって、車両と駐車位置に取り付けたマーカを認識することで取得する。

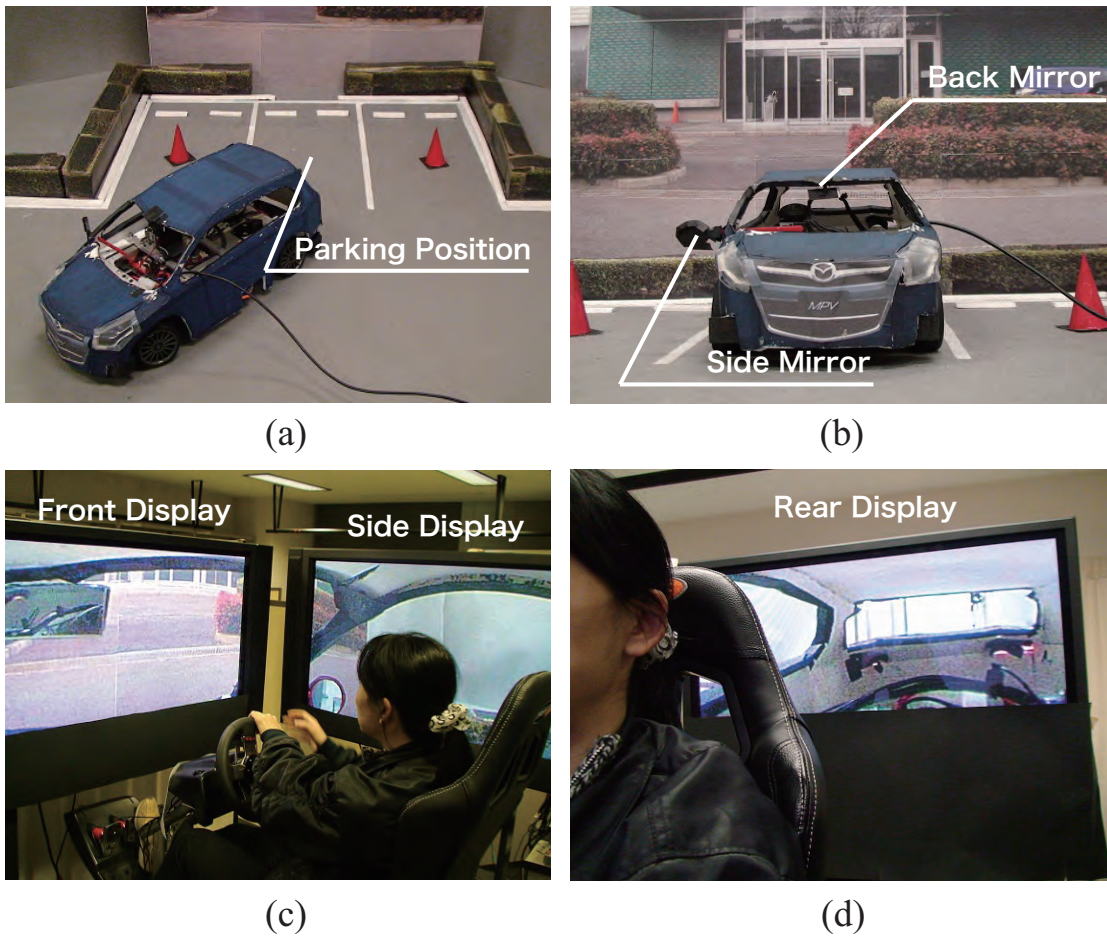


図 2.9: Scaled car simulator. (a) Bird's eye view, (b) Front view, (c) Front view from the driving seat, and (d) Rear view from the driving seat.

2.5.2 経路計画手法

提案する運転支援システムの実現には、切返し位置と駐車目標位置を結ぶ経路の動的な生成と、それに沿って車両を制御するための手法が必要となる。そこで、任意の切返し位置からの動的な経路計画を実現するため、図 2.10 に示すように切返し位置での車両の姿勢、及び目標の位置・姿勢から Bezier 曲線を作成し、車両制御のための目標経路とする。駐車軌道の生成手法としては緩和曲線の一種であるクロソイド曲線が一般に知られているが [46]、任意の切返し位置・角度から駐車目標位置までのクロソイド曲線を求める逆問題は計算コストが高く、実時間での

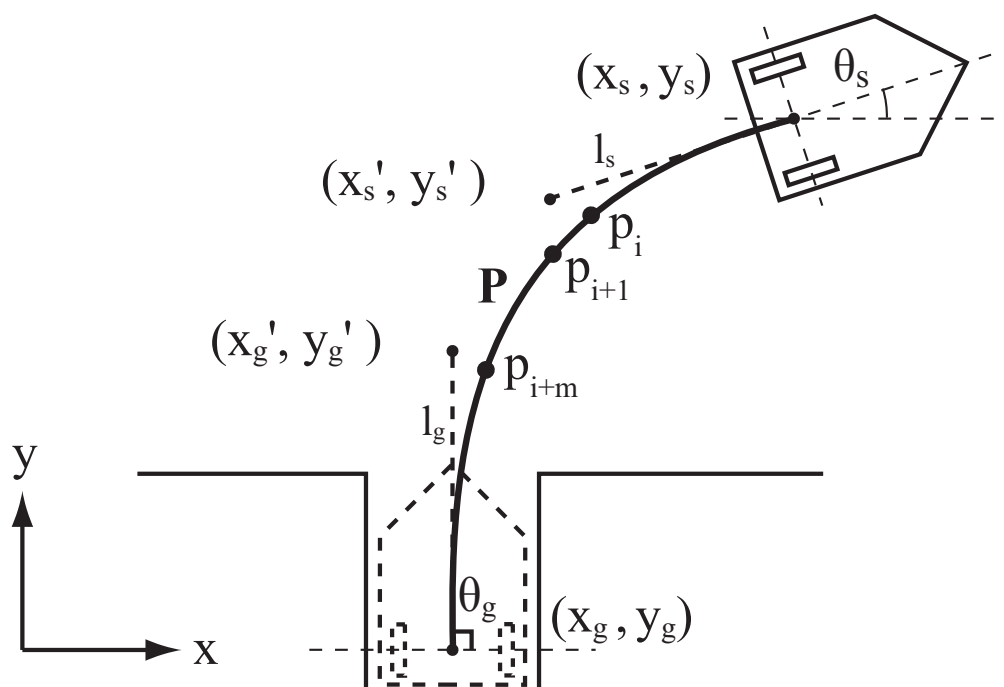


図 2.10: Path-planning algorithm based on Bezier curve.

逐次支援には向かないという問題がある。一方、Bezier 曲線では任意の二点間を結ぶ曲線を端点での曲率を制御しながら一括で高速に計算可能なため実時間性が高く、運転者への切返し位置の指示や、切返し位置で計算のため長時間停車する必要なく支援を開始出来る。また、ステアリング制御や最適入力 of 計算過程で急激なハンドル操作（急回転・停止等）を抑制することで、クロソイド曲線を用いることなく違和感の少ない操作支援を行うことが可能となる。

切返し地点における車両の位置・姿勢を (x_s, y_s, θ_s) 、駐車目標位置・姿勢を (x_g, y_g, θ_g) とする。次に、それぞれの端点における曲率を制御する接線の長さを l_s, l_g として制御点を

$$(x'_s, y'_s) = (l_s \cos(\theta_s + \pi), l_s \sin(\theta_s + \pi)) \quad (2.9)$$

$$(x'_g, y'_g) = (l_g \cos \theta_g, l_g \sin \theta_g) \quad (2.10)$$

と定める. 曲線の分割数を n とすると, 目標経路 $P = \{p_i(x_i, y_i) | i = 0 \cdots n\}$ が次のように求まる.

$$t = i/n \quad (0 \leq t \leq 1) \quad \tau = (1 - t) \quad (2.11)$$

$$x_i = \tau^3 x_s + 3t\tau^2 x'_s + 3t^2 \tau x'_g + t^3 x_g \quad (2.12)$$

$$y_i = \tau^3 y_s + 3t\tau^2 y'_s + 3t^2 \tau y'_g + t^3 y_g \quad (2.13)$$

$$\theta_i = \arctan \left(\frac{y_i - y_{i+1}}{x_i - x_{i+1}} \right) \quad (2.14)$$

また, 求めた経路 P の長さ L , 最大曲率 κ は以下の通りとなる.

$$L = \sum_{i=0}^{n-1} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \quad (2.15)$$

$$\kappa = \max_i \left| \frac{\dot{x}_i \ddot{y}_i - \dot{y}_i \ddot{x}_i}{(\dot{x}_i^2 + \dot{y}_i^2)^{\frac{3}{2}}} \right| \quad (2.16)$$

ここで, \dot{x}_i は x_i の 1 階微分, \ddot{x}_i は 2 階微分を表す. 今回, 一定の範囲内の l_s, l_g に対して上記の計算を行い, 最大曲率の逆数 $1/\kappa$ がスケールカーの最小旋回半径以上であり, かつ経路長 L が最短となる経路を最適経路 \hat{P} とし, この経路上を移動する点に対する目標位置追従問題として自動駐車制御を実装した.

また車両の幾何学的な制約上, 上記の手法によって現実的な駐車軌道を計算することが不可能な切返し位置も存在する. その場合は, ステアリングに実装された LED によって運転者に切返し位置が不適切であることを教示するとともに, その後の運転操作への介入は行わず運転者が単独で駐車することとする.

2.5.3 力覚呈示算出手法

力覚呈示による操作支援が視覚や聴覚への情報呈示による操作支援と比較して優位な特徴として、操作すべき量・タイミングを運転者が直感的に知覚し易いことは先に述べた通りである。しかしながら、力覚を適切なタイミングで運転者に呈示するためには、コントローラへの入力からステアリングに内蔵されたアクチュエータによるトルクが発生するまでの時間遅れが問題となる。そこで、現在の車両の状態から一定時間先の車両状態を予測し、その時点でのステアリング角を目標角度とするステアリングの制御を実装することで制御時間遅れによる操作教示のタイミングの遅れを解消する。また、提案する支援手法ではシステムによるステアリングの制御中に運転者による自発的な操作が外乱として入力されることが想定されることから、今回ステアリング呈示力覚の算出には逐次最適化型の制御手法である Receding Horizon 制御を用いる [57] [58]。これは、通常最適制御手法が制御入力を事前に求めるのに対し、現時点から有限時間先までの最適制御を連続的に解くことで制御を行う手法である。このため、外乱に対し頑健でかつ計算量を削減することで実時間制御に適した手法となっている。ここでは、前節で作成した経路に沿った Receding Horizon 制御による最適入力の計算過程で得られる車両状態の将来予測値に基づいて、運転者に呈示する力覚量を求めるアルゴリズムを考案した。

Receding Horizon 制御による制御入力の算出にあたり、車両の状態変数を $\mathbf{x} = \{x, y, \theta, \delta\}$ 、入力を $\mathbf{u} = \{u_1, u_2\}$ とし、車両の状態方程式には以下のような幾何モデルを用いる。

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}, t) = \begin{bmatrix} u_1 \cos \theta \\ u_1 \sin \theta \\ u_1 \frac{\tan \delta}{D} \\ u_2 \end{bmatrix} \quad (2.17)$$

ここで、 (x, y) は車両の2次元座標 [m]、 θ は車両の姿勢 (向き) [rad]、 δ はステアリ

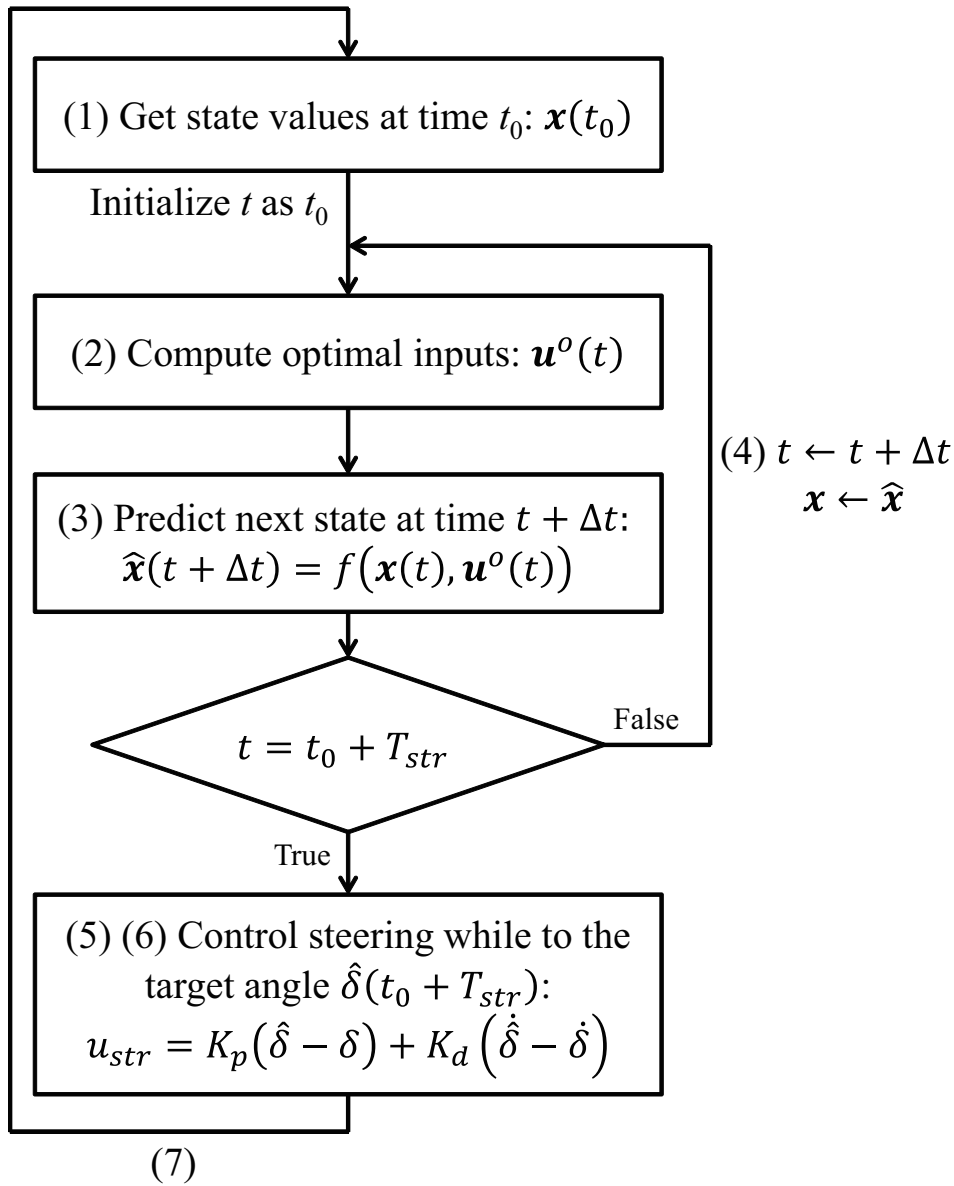


図 2.11: A flow chart of haptic instruction calculation.

ング角 [rad], u_1 は車両に対する速度入力 [m/sec], u_2 は車体の正面方向に対するステアリング角速度 [rad/sec], D はホイールベース距離 [m] を表す。

また, 車両の状態変数の更新および制御周期は 100[ms] とし, 呈示力の計算は各離散時刻毎に図 2.11 に示す以下の手順で行う。

1. 時刻 t_0 における車両の状態を $\mathbf{x}(t_0)$ とし, $\mathbf{x}(t_0)$ に対する経路 P 上の最近傍点

が \mathbf{p}_i であるとき, m を正の整数として $j = i + m$ で表される点 \mathbf{p}_j を短期的な目標位置・姿勢として定める.

2. \mathbf{p}_j に対し, $t_0 \leq t \leq t_0 + T_f$ の時間区間で最適制御問題を解き, 最適入力系列 $U^o = \{\mathbf{u}^o(t) | t_0 \leq t \leq t_0 + T_f\}$ を求める. なお評価関数には車両の現在位置・姿勢と目標位置・姿勢との誤差, 及び制御入力を最小化するために以下の式を用いる.

$$\begin{aligned} J &= \int_0^{T_f} (\mathbf{x}_g - \mathbf{x}(t))^T Q (\mathbf{x}_g - \mathbf{x}(t)) dt \\ &+ \int_0^{T_f} \mathbf{u}(t)^T R \mathbf{u}(t) dt \\ &+ (\mathbf{x}_g - \mathbf{x}(T_f))^T S (\mathbf{x}_g - \mathbf{x}(T_f)) \end{aligned} \quad (2.18)$$

\mathbf{x}_g は駐車目標位置・姿勢を表すベクトルであり, Q, R, S は重み行列を表す.

3. $\mathbf{u}^o(t_0)$ と車両の状態方程式 f から時刻 $t_0 + \Delta t$ における車両の状態変数 $\hat{\mathbf{x}}(t_0 + \Delta t)$ を推定する. ここでの数値計算には4次のルンゲ・クッタ法を用いる.
4. $\mathbf{x} \leftarrow \hat{\mathbf{x}}, t_0 \leftarrow t_0 + \Delta t$ として, 上記1~3を繰り返すことで, 一定時間先までの車両状態の予測値 $\hat{X} = \{\hat{\mathbf{x}}(t) | t_0 \leq t \leq t_0 + T\}$ を得る.
5. 入力インタフェースのステップ入力に対する時定数を T_{str} とし, 時刻 $t_0 + T_{str}$ におけるステアリング角の推定値 $\hat{\delta}(t_0 + T_{str})$ をステアリング制御の目標角度に設定する. なお, 今回設定したPDゲインによる目標ステアリング角へのステップ応答の時定数 T_{str} は約500[ms]である.
6. 算出した目標ステアリング角に対し, PD制御による角度制御を行う. 制御入力 u_{str} は, 現在のステアリング角度, 角速度に基づき以下のように求める.

$$u_{str} = K_p(\hat{\delta} - \delta) + K_d(\dot{\hat{\delta}} - \dot{\delta}) \quad (2.19)$$

ここで, u_{str} は今回使用するステアリング型入力インタフェースへのトルク指令入力を表し, K_p, K_d はそれぞれ比例ゲイン, 微分ゲインを表す. 目標角速度 $\dot{\hat{\delta}}$ は常に0とし, 目標角度との誤差に応じて $0.2[N] \sim 0.6[N]$ の範囲で線形

に与えられるように K_p, K_d を調節したところ、運転者がステアリングから手を離した無負荷状態で任意の切返し地点から自動駐車が可能になることを確認した。

7. 上記の処理を約 10[Hz] で繰り返し行う。

2.6 支援効果検証実験

2.6.1 実験手順

提案する支援システムによる支援効果の検証するため、前述したスケールカー・シミュレータを用いて実験を行う。実験課題は第3章で述べた駐車行動計測実験と同様とし、参加する被験者 10 名 (男 8 名, 女 2 名, 運転歴 12~30 年) を事前に運転歴が均等になるよう二つのグループに振り分け、それぞれ A 群, B 群とした。なお、今回の被験者の内 8 名は 3 章で述べた実車実験にも参加しているが、数ヶ月の時間的な開きがあるため実質的な影響は無いとする。駐車向きは右バック駐車のみとし、被験者ごとに 5 回の駐車試行を 1 セットとして 3 セット (15 回) 行うよう指示し、その様子を観察する。また、スケールカーが走行する実験環境は実験室内の様子が見えないよう床と同じ単一色の壁で囲われており、壁には目印となるようなオブジェクトは取り付けず、目標の駐車スペースは左右の他のスペースに赤いロードコーンを配置することで運転者に視覚的に提示する。シミュレータの座席高さ、およびサイドミラー・バックミラーの角度は実験開始前に被験者ごとに入念に調節を行い、シミュレータの運転席から窓・ミラーを通して駐車位置を示す白線やロードコーンが十分視認出来ることを予め確認した。

実験時における提案手法による運転支援の実施条件を表 2.4 に示す。支援を実施する対象は B 群の第二セットのみとし、A 群の全試行および B 群の第一、第三セットでは支援を行わない被験者単独での駐車とした。これにより、各被験者群の第一セットの駐車成績に対する第三セットの駐車成績の改善度合いを技量の向上幅と見做し、両群の比較を行うことで運転支援システムによる介入の有無が技

表 2.4: Conditions of experiment

	1 st set (1~5 trials)	2 nd set (6~10 trials)	3 rd set (11~15 trials)
Group A	non-assisted	non-assisted	non-assisted
Group B	non-assisted	assisted	non-assisted

量向上へ与える影響を明らかにする。なお、今回の実験における呈示力の強度や呈示するタイミング等の支援パラメータは、実験開始前に経験的に定めたものを用い各被験者間で一定とした。また、以降の解析には実車実験に参加した10名の内、実験プロトコルの確認・修正のために行った予備実験に参加し、その過程で実験環境に変更のあったA群の被験者1名と、後述する定量化した駐車成績が第一セット開始時の段階で他の被験者よりも有意に高かった ($p < 0.01$, U検定) B群の被験者1名を除いた、計8名分の計測データを用いることとする。

2.6.2 駐車成績の解析結果

実験結果の解析にあたり、まず駐車行動計測実験と同様に駐車成績の定量化を行う。ここでは、駐車位置での並進・回転誤差および、切返し位置からの所要時間を、全被験者の第一、第三セットでの計測値の平均値と標準偏差を用いてそれぞれ正規化し、足し合わせたものを駐車成績として解析に用いる。

両被験者群の第一、第三セットにおける駐車成績の平均と標準偏差を比較したものを図 2.12 に示す。この図から、第一セットでは両群の駐車成績に有意な差は見られなかったにも関わらず、第三セットでは支援を行ったB群の被験者の駐車成績が、支援を行わなかったA群の被験者の駐車成績に比べて向上していることが見て取れ、t検定を用いた有意差検定でも危険率1%未満で有意であることが示された。

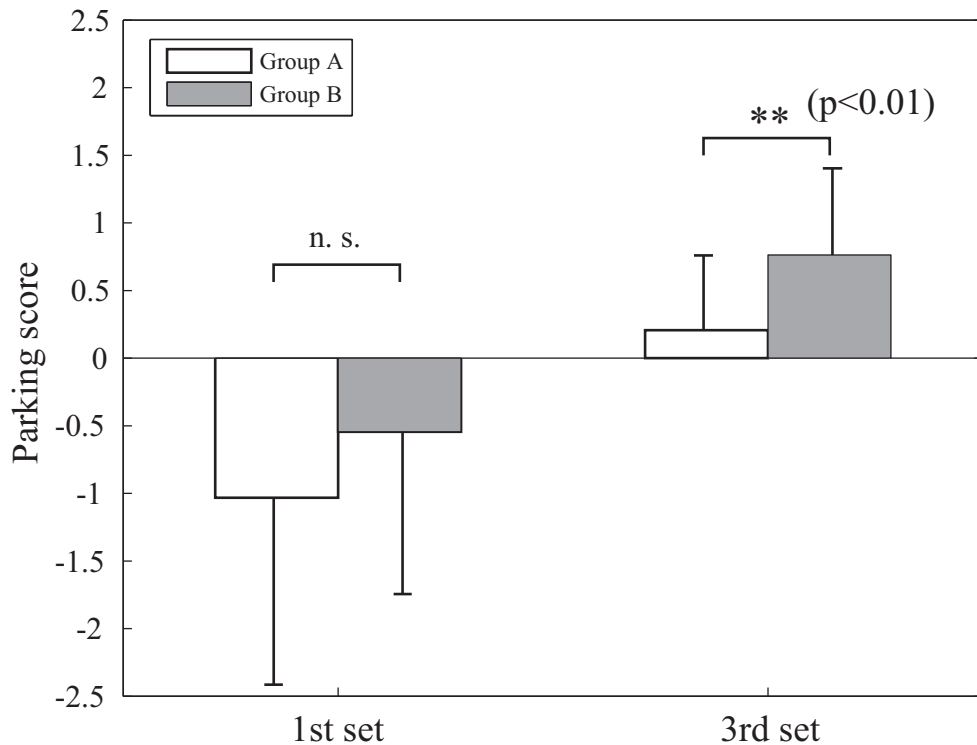


図 2.12: Comparison of averaged parking scores of the groups.

2.6.3 切返し位置と駐車成否の関係

今回実験タスクとして設定した後ろ向き駐車運転行動は、初期位置から切返し位置までの運転行動と、切返し位置から駐車位置までの運転行動という二つの運転課題から構成される。ここで、実装した力覚呈示による操作支援は切返し位置以降の運転操作を教示の対象としているが、車両の幾何学的特性と切返し回数の制限のため、駐車成否および後退時における運転操作の難易度は原理的に切返し位置に依存する。そのため、支援対象である切返し位置以降の運転操作量・及びタイミングに関する提案手法の支援効果を検証するために、切返し位置が適切であった場合とそうでなかった場合を区別した上で駐車成績を評価する。ここでは、現在の切返し位置から目標位置に至る経路が生成出来る位置を適切な切返し

表 2.5: Relationship between the direction-change point (DCP) and the rate of successful parking

	Group A		Group B	
	1 st	3 rd	1 st	3 rd
(1) Reached at the appropriate DCP	0.6	0.68	0.45	0.67
(2) Parking succeeded	0.1	0.53	0.1	0.65
(3) Parking succeeded from the appropriate DCP	0.17	0.69	0.11	0.71
(4) Parking succeeded from the inappropriate DCP	0	0.17	0.09	0.5

位置と定義し、切返し位置の妥当性と駐車成功率の関係について分析した結果を表6に示す。なお、ここでは並進誤差が0.1 [m] 未満かつ回転誤差が0.1 [rad] 未満の試行を駐車成功と見做し解析を行った。この表から、(1) 適切な切返し位置に到達した割合、(2) 駐車成功した割合、および(3) 適切な切返し位置から駐車成功した割合は、支援の有無に関わらず両群ともに上昇していることが見て取れる。その一方で、(4) 適切な切返し位置外から駐車成功した割合は、第三セットにおいてA群が17%であるのに対して、支援を行ったB群では50%と大きく差が見られる結果となった。

2.6.4 切返し位置の分散に関する解析結果

また、適切な切返し位置の学習過程に対する提案手法の影響について検証するため、被験者ごとの切返し位置の分布とその変化についても解析を行った。図2.13は、両グループから1名ずつ抽出した被験者の駐車軌道及び切返し位置の一例を表す。この図から分かるように、両群の被験者の駐車軌道を比較すると特に支援を行った被験者群の第三セットにおける切返し位置のバラつきが収束していることが見て取れる。そこで、同一被験者の切返し位置を表す座標点の集合に対し主成分分析を行い、その第一主成分を切返し位置の最大分散値として定量化し、両

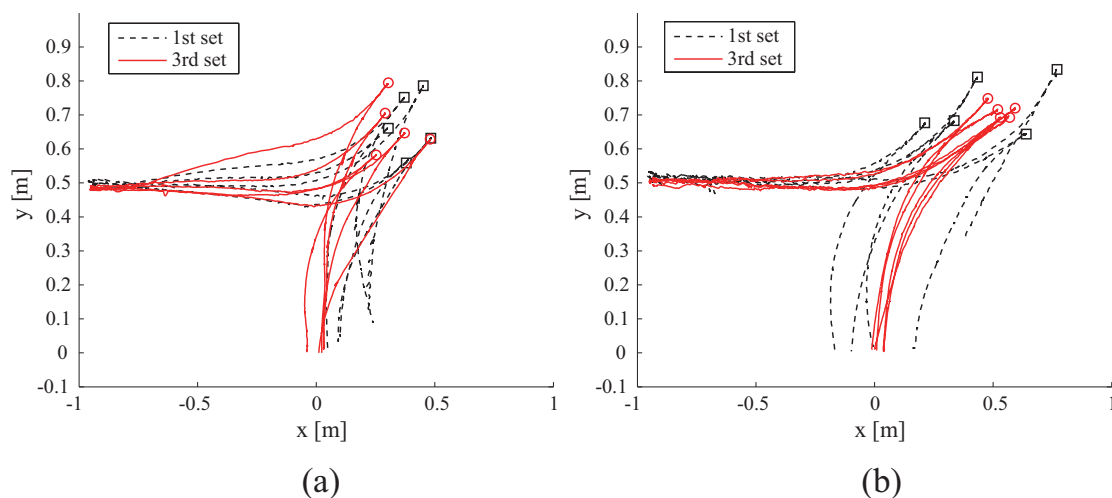


図 2.13: Examples of parking trajectories. (a) subject in group A, (b) subject in group B.

群の平均分散値の変化を比較したものが図 2.14 である。この図から支援を行わなかった被験者の切返し位置の分散は第一セットと第三セットの間に統計的な変化が無かったのに対し、支援を行った被験者群では第三セットにおいて切返し位置のバラつき (分散) が有意に減少していることから、第三セットにおいてほぼ同じ切返し位置から繰り返し駐車を行っていることが分かる ($p < 0.05$, U 検定)。なお、駐車目標位置への経路が生成出来る切返し位置はある程度の範囲に広がりを持つため、表 2.5 で示した適切な切返し位置に到達した割合が等しいという結果と、切返し位置の分散に差が見られたことは矛盾しない。

2.6.5 考察

以上の解析結果より得られる考察として、まず駐車位置の並進・回転誤差および駐車に要した時間から求める総合的な指標である駐車成績の変化を示した図 2.12 から、実験開始時点での両群の被験者の駐車運転技量に有意な違いは無かったと見做すことが出来る。それが第三セットの時点では両群の駐車成績に統計的に有意な差が見られたことから、これは B 群の第二セットで実施した提案手法による

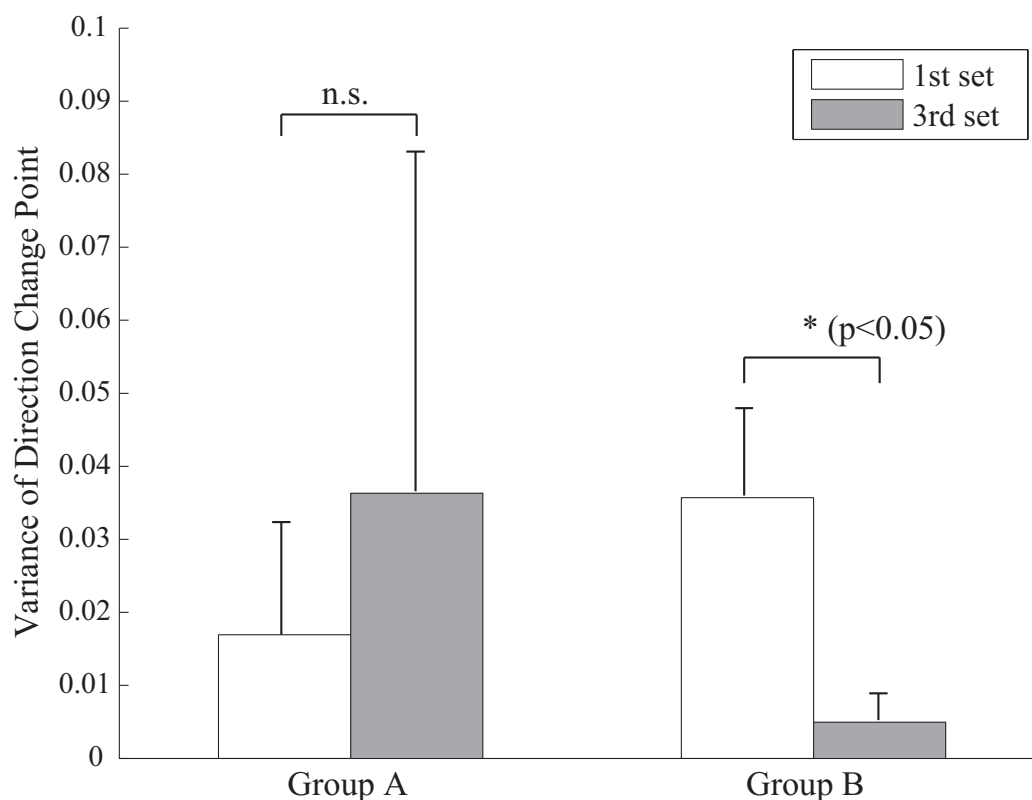


図 2.14: Absolute variance of the direction change point.

支援の影響であると考えられ、今回開発した運転支援システムによる駐車技量向上の試みは一定の効果があったと判断出来る。

また、切返し位置と駐車成否の関係については、表 6 より結果として駐車が成功した割合は両群の第一セットで等しかったにも関わらず、特に (4) 適切な切返し位置外から駐車が成功した割合に有意な差が確認出来た。このことから、上記の駐車成績の違いは主に切返し位置以降の運転操作量・およびタイミングに関する運転技量の向上に依るものであったと考えられる。つまり、切返し位置が適切であっても常に駐車が成功するとは限らず、一方で、切返し位置が厳密に適切でなくとも妥当な位置・姿勢での駐車は可能であることから、切返し位置以降の運転技量の習熟度合いが最終的な駐車成績に寄与する割合は大きいと言える。よって、

B群の被験者の学習過程が提案する操作支援によって促進されたことが、第三セットにおいて適切な切返し位置から駐車を行った割合が同程度にも関わらず、駐車成功率及び駐車成績に差が見られた理由であると考えられる。

一方、適切な切返し位置に到達した割合は両群とも第三セットで同程度に収束していることから、切返し位置の選定に関する提案手法の直接的な効果は確認出来なかった。これは、今回切返し位置に関しては可能な限り運転者の任意とし、位置の選定に対する直接的な支援は行わなかったことから妥当な帰結であると言える。しかしながら、図 2.14 に示すように支援を行った被験者群の切返し位置の分散は、対照群に比べて第三セットにおいて有意に減少する傾向が見られた。これは、学習過程において A 群の被験者は第一セットの時点で既に高い割合で適切な切返し位置を選んでいたにも関わらず (60%)、駐車失敗の原因がその後の運転操作ではなく切返し位置にあると考え、試行の大半を切返し位置の探索に費やしたことが A 群で切返し位置が収束せずむしろ平均的には増加した原因と考えられる。一方、B 群の被験者はシステムの支援によって試行の比較的早い段階で駐車成功を経験し、かつそれを再現するための切返し位置、及びそこからのステアリング操作量とタイミングを力覚呈示によって同時に直感的に学習出来たため、切返し位置が収束するとともに切返し位置以降の操作技量が向上したことで適切な切返し位置外からの駐車精度も改善したと思われる。また、第二セット中で不適切な切返し位置に対しては LED によるフィードバックが与えられたことも、切返し位置の探索行動を早期に完了させるのに有効であったと考えられる。

また、今回スケールカーが走行した実験環境は駐車位置を示す白線や左右の駐車スペースに置かれたロードコーン以外に運転者が自己位置を推定するための目印は設置していない。従って被験者は初期位置から切返し位置までの移動中、前方および右側の窓越しに見える目標駐車位置を確認しながら切返し位置を定め、後退中は後方の窓やサイドミラーから推定した目標位置との相対位置・姿勢に基づき操作量・タイミングを記憶・学習したと考えられる。このことから、本実験で習得した後ろ向き駐車運転の運転技量は、周囲の景色や建物が異なる運転環境にも応用可能であると考えられる。

2.7 第3章まとめ

本章では，人間機械系の物理的な相互作用に基づく行動学習支援の事例研究として取り組んだ，駐車運転技量育成システムの開発について述べた．ここでは，ステアリングへの力覚呈示によって人の駐車運転操作の学習を支援するシステムを提案・開発するとともに，被験者実験によってその有効性を明らかにした．このとき，ステアリングへの力覚呈示に用いるアクチュエータの応答特性による時間遅れを補償するアルゴリズムや，熟練者の運転特性を再現する評価関数を用いることで，随伴性と一貫性のあるシステムから人への物理的な教示方法を実現した．

第3章 情動的相互作用に基づく行動学習支援

3.1 事例研究の位置づけと概要

次に、人と機械の情動的相互作用に基づく行動学習支援に関する事例研究として、教示者によるロボットの行動学習支援について述べる。これは、行動学習を行うロボットに対して人が実時間で主観的な評価を与えることで学習を支援するための枠組みである。ここでは、人の教示に関する随伴性・一貫性に基づき教示者の教示意図を推定し評価関数を動的に更新する手法を提案し、それによって評価関数の設計に関する既存の行動学習手法の問題を解決するとともに、人による機械システムの行動学習支援についての知見を得ることを目的として行った。

3.2 背景

3.2.1 一般的な行動学習手法の問題点

ロボットによる行動学習は一般に、行動選択→評価→パラメータの調節という一連の流れを繰り返すことで行う。その際、ロボットが取り得る全ての状態に対して事前に正しく評価関数を設計できることが理想的であるが、現実的には困難である。そこで、ロボットによる自律的な行動獲得を実現する代表的な手法として強化学習がある [2]。強化学習は、時刻 t における状態 $s(t)$ での状態価値 $V(s)$ を環境から得られる報酬 $r(t)$ に基づいて更新し、最終的に得られる報酬の期待値を最大化するような行動系列を試行錯誤的に求める学習手法である。そのため、強化

学習は実環境での行動学習に適しており、様々な手法が提案されてきた。例えば、TD(λ)法は報酬遅延に対応した強化学習の代表的な手法であり、学習エージェントがある時点で報酬を得ると、 λ ステップ前の時点までの状態に対し、報酬を割引きながら状態価値の更新を行う。これより λ を大きくすることで、エージェントはより将来の報酬も考慮した学習を行うことが可能となる。また同様に、報酬遅延に頑健とされる経験強化型の学習手法 (Profit-sharing) がある [59]。この手法は、エージェントが報酬を得た時点でそれまでの行動系列を一括で強化するため、一般的に TD(λ)法よりも早く学習が収束することが知られている。

しかしながら、これらの手法はいずれも学習開始から一定時間内に何らかの報酬が得られることを前提としている。一般的な強化学習では、エージェントは初期状態からランダムに探索を行い、初めて報酬を得た時点で学習を開始する。そのため、初期状態から探索可能な範囲で有限時間内に報酬を得られる確率が低い場合、長時間報酬に辿り着くことがないまま学習が進行しない状態が続く可能性がある。そのような事態を避けるため、通常設計者はエージェントが学習可能な報酬関数の調節を入念に行う。しかし、報酬関数は原則的にタスク依存であり、エージェントによって学習可能であると共に収束が保証される報酬関数の作成のための統一的な方法論は未だ確立されていない。そのため現状では、使用する学習手法やタスクに応じて報酬関数を人間が経験的に設定することが必要不可欠である。特に実環境中で動作するエージェントが複雑な学習を行う際には、常に適切な報酬関数を事前に設計できるとは限らず、使用する学習アルゴリズムにとって最適な報酬関数を求めるために試行錯誤を繰り返しながら設計を行う必要がある。

3.2.2 関連研究

そこで近年、模倣学習 (Imitation learning) や例示による動作生成 (Programming by demonstration) など、教示者とのインタラクションによって複雑かつ多様な行動を学習するための手法が注目されている [10] [7]。しかしながら模倣や見まねによる学習は、教示側と模倣側がある程度類似した身体構造や自由度を持っている

ことを前提としており適用範囲に限られる。また、例示による学習では教示者による実演を必要とするため人による実演が困難なタスクの学習には向かず、実演のために必要なインタフェースもタスクやプラットフォームによって異なるといった問題がある。さらに、上記の手法での行動獲得は周囲の環境との相互作用を伴わないため、獲得した動作の目的やその動作が外界に対して及ぼす作用などはエージェントによって学習されない。そのため、獲得した動作を用いて環境中で明確な目的を持った意味のある行動を学習するためには、依然として強化学習のような経験強化型の行動学習の枠組みが必要であるといえる。

Lockerdらは、強化学習を実装したエージェントに対して人が対話的に報酬を与える Socially-Guided Machine Learning を提案し、人がエージェントの学習を監督することでより効率的な行動学習を行えることを示した [13]。また前田らは、動物行動学やトレーニングに広く用いられている Shaping の概念をエージェントの行動学習支援に応用し、教示者による段階的な行動の強化を行う分化強化型の学習手法によって複雑な行動系列を段階的に学習できることを示した [60]。しかしながら、これらの先行研究ではどちらも人による報酬を通常の報酬の代替として用いており、人から与えられた個々の教示の背後にある大局的な意図や方策を推定し、評価関数に反映させるための仕組みではない。また、エージェントの行動とそれによる状態遷移は離散的なマルコフ過程を前提としており、実環境で動作するロボットの行動学習支援に直接的に適用することは出来ない。

3.2.3 目的

この問題に対し本研究では、人が教示者として機械系の学習を監督し対話的にフィードバックを与えることで、評価関数を動的に修正する行動学習支援手法を提案する。ここでは、図 3.1 に示すように人が機械系の学習主体であるエージェントの振る舞いを観察しながら主観的な評価を教示として実時間で与え、学習エージェントは与えられた教示と観測される周囲の環境や状態の変化を関連付けることで人の教示意図を推定し評価関数を動的に修正するという人間機械系の学習モ

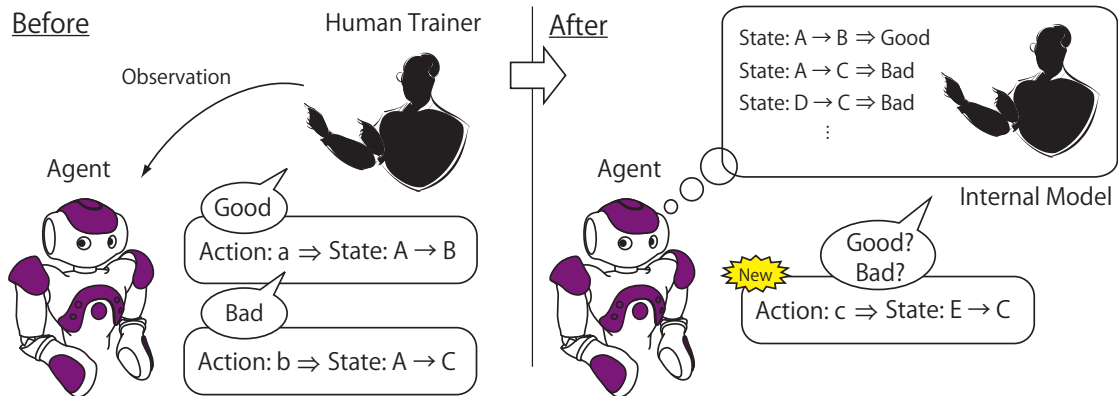


図 3.1: Behavior learning assistance based on human subjective cues.

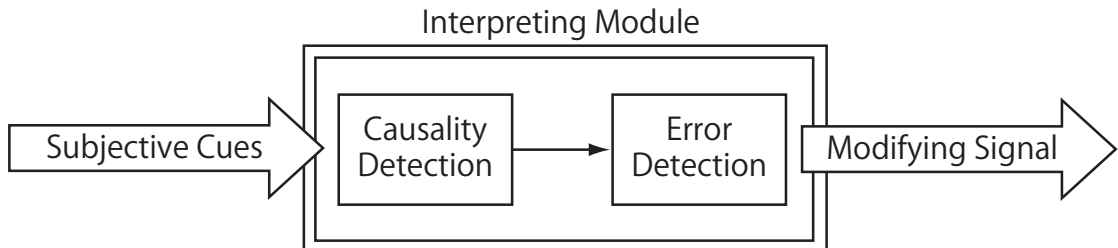


図 3.2: Causality detection and Error detection.

デルを提案する。これは、従来独立して行われていたエージェントによる行動学習と、人による事前の報酬設計の作業を並行してかつオンラインで行うための手法であり、これによってエージェントの行動学習を支援すると共に設計者による事前の報酬設計の負担を軽減することもできると考えられる。

3.3 教示者によるロボットの行動学習支援

3.3.1 主観評価の導入

教示者がエージェントに与えるフィードバックには、「良い/悪い」という主観的な2値の評価を用いる。これは、正負の評価を表す基準として単純かつ幅広いタスクに対して意味を持つ指標であり、原理的には人が「良い」もしくは「悪い」と判

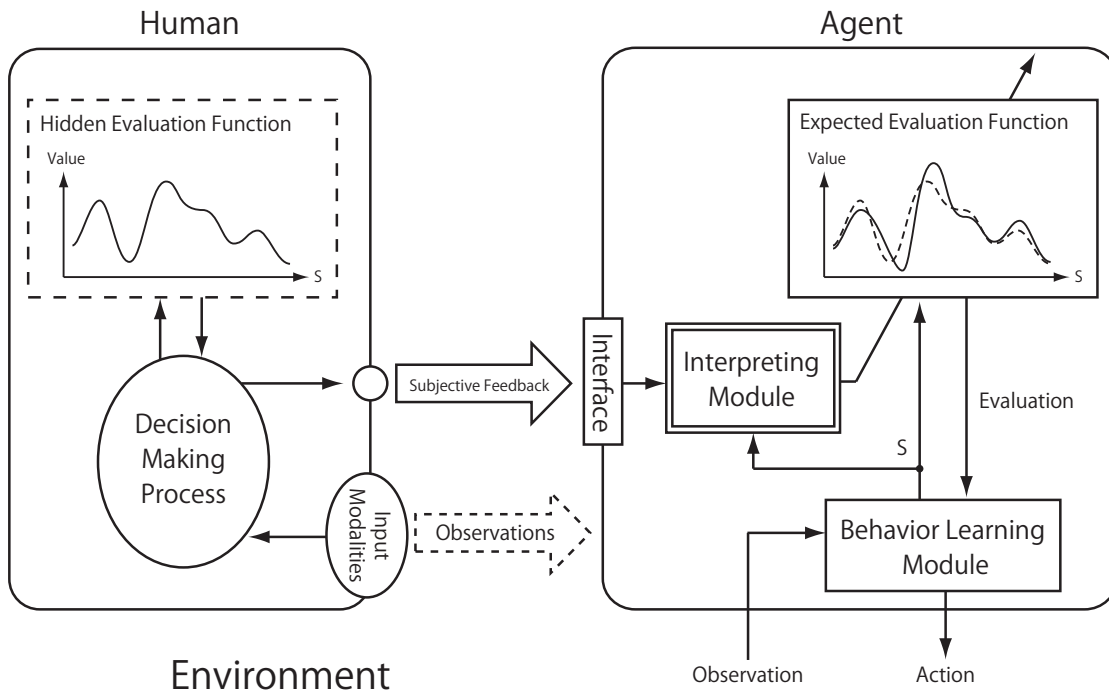


図 3.3: Behavior learning assistance based on human subjective cues.

断出来るタスクに対して教示の種類やインタラクションの枠組みを変更することなく提案手法を適用することが可能となる。また、内容を細かく指定する必要が無いため素早く直観的な判断が容易であることから、実時間でのインタラクションに適した指標であるといえる。ここで、単純な二種類の教示に基づいて多次元の状態空間上に定義される評価関数を更新するためには、その教示が与えられた時のエージェントの行動や状態から人の教示意図を推定して評価関数に反映させる必要がある。そのとき、人の認知判断過程で生じる時間遅れや誤認知や誤判断により間違った教示が与えられる可能性を考慮し、図 3.2 に示すような人の教示の随伴性と一貫性に基づく起因の特定と誤り検出によって、人からの教示を評価関数の学習に利用可能な情報として解釈するメカニズムを、従来の行動学習モジュールとは独立して実装する (図 3.3)。

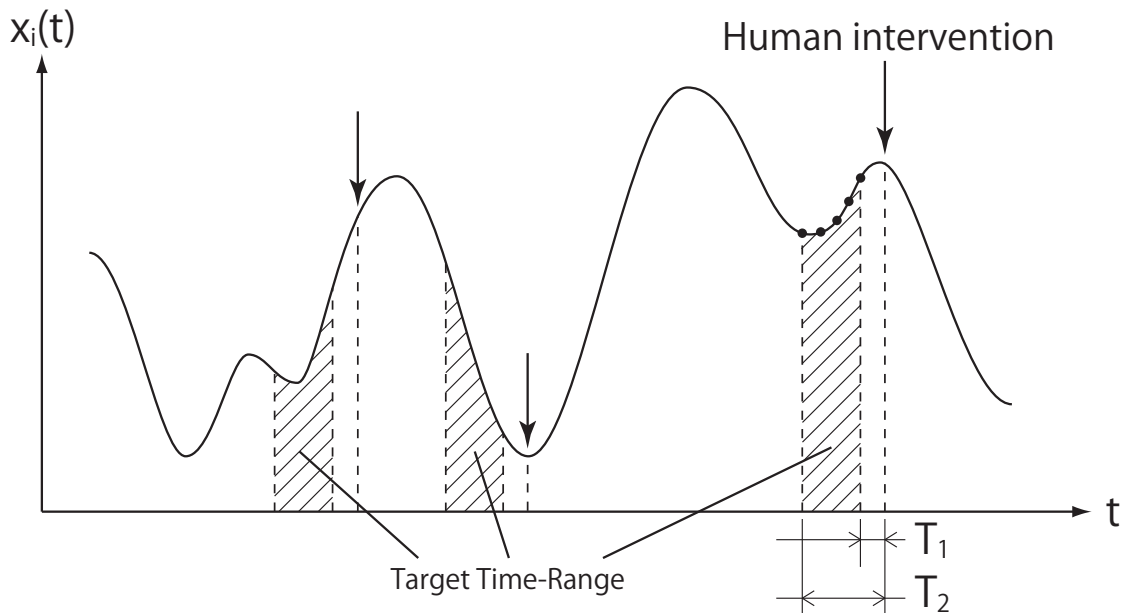


図 3.4: Time-consistency based human intention estimation.

3.3.2 人の教示の随伴性・一貫性に基づく教示意图推定

始めに教示の時間遅れを測定する実験を複数人の被験者(20代男性4名)による協力の下行った。ここでは、エージェントによる連続した動作の中から特定の動作の連なりを認識した時点で入力を行い、そのときの時間遅れを測定する。その結果、いずれの被験者においても対象動作の観測から1秒以内に教示動作を行っており、個人差はあるが平均的な時間遅れはどの被験者も同程度であることが分かった。そこで、図3.4に示すように T_1, T_2 という時間遅れ定数を定義し、教示が与えられた際にこれらの時定数に基づき時間的に限定された範囲内で教示の対象となる行動の候補を時間的に絞り込むために用いる。これによって、学習中に教示者から評価が与えられたときエージェントは過去の全ての状態を参照する必要が無く、計算コストを削減することが出来る。また、人による主観的評価には一貫性がある、すなわち同様の状況においては同様の教示が与えられるとの仮定から、過去の教示の履歴の参照し教示の一貫性を検証することで、誤教示による外乱も抑制することが出来る。

3.3.3 行動学習支援のアプローチ

図 3.5 に、評価関数の動的更新によるエージェントの行動学習支援のアプローチを示す。ここでは、機械系の行動学習手法として広く用いられている強化学習に適用することとする [2]。図の上段は、強化学習の枠組みにおける評価関数に相当する報酬関数が目標状態付近でのみ定義され、エージェントが初期状態から探索可能な範囲内で報酬を得られず学習が困難な状況を表している。このようなケースに対し、教示者がエージェントの学習状況に応じた適切な教示を行うことで、図の下段のように初期状態から目標とする状態まで連続的で滑らかに増加する報酬関数を作成することができる。これにより、様々な問題に対して予め最適な報酬関数を設計する必要なく強化学習による行動学習メカニズムを実装できるとともに、環境やタスクの変化にもその都度教示者が介入することで柔軟に対応することが出来る。

なお、本提案手法での教示者によるエージェントの行動学習支援は、初期状態から目標とする状態まで連続的かつ滑らかな状態遷移による行動学習タスクを対象としている。そのため、タスクの分解及び学習器の階層化を考慮しておらず、最終的な目標を達成するために幾つか特定の状態を経由する必要があるタスク、いわゆるサブゴールのあるタスクに対しては提案手法を直接適用することは想定していない。しかしながら、一般にサブゴールのあるタスクは目標状態に向かって連続的に報酬関数を定義できる区間を構成要素(サブタスク)として分解し、学習器のモジュール化や階層構造の導入によってサブタスクを順次切り替えながら学習を行うことで達成される。そのため、目標状態まで滑らかに報酬関数を形成しエージェントの学習を支援する提案手法は、サブゴールのある問題に対しても各サブタスク内での行動学習支援として適用可能であると考えられる。

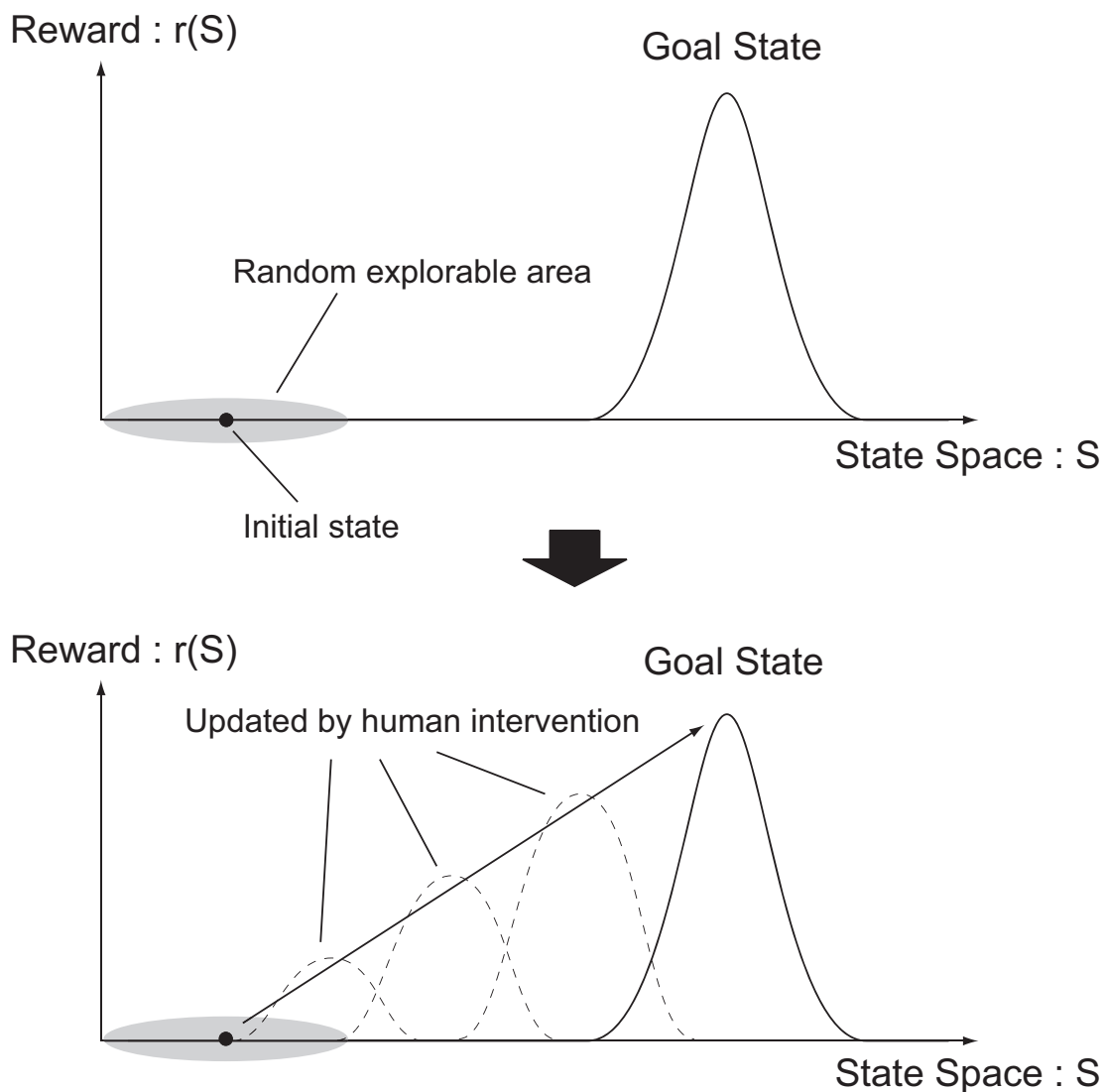


図 3.5: Approach of behavior learning assistance for reinforcement learning agent.

3.4 実験環境

3.4.1 実験課題

ここで、次節以降の検証実験で用いる実験環境および行動学習課題について説明を行う。実験タスクは、図 3.6 に示すような 6 自由度のロボットアームの手先に取り付けた 1 リンクの振り子の振り上げと倒立維持とし、エージェントは振り子

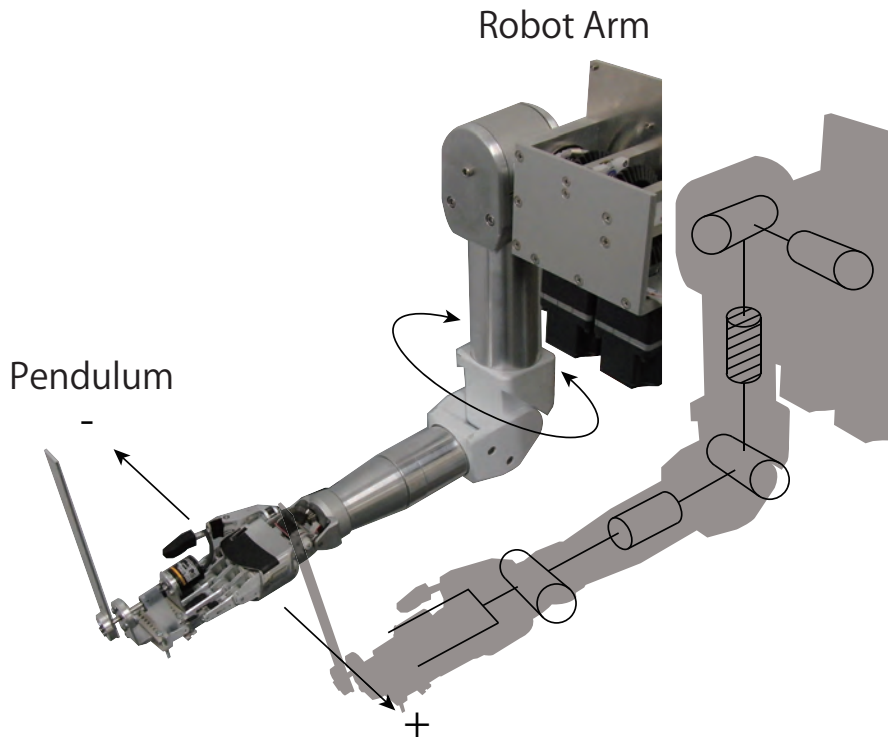


図 3.6: 6 DOF robot arm.

の角度，角速度からなる二次元の状態空間における図中斜線で示した上腕の関節制御入力の学習を行うこととする．本研究では，図 3.6 に示す実機および，図 3.7 に示すようなシミュレーション環境の両方を用いて検証実験を行った．

3.4.2 教示入力インターフェース

提案手法を実ロボットの行動学習支援に適用する際，人がロボットに対して教示を与えるためのインターフェースが必要となる．本研究では，サイズや形状の異なる様々な実ロボットに対して直観的かつ簡便に主観フィードバックを行うためのインターフェースを開発した．本インターフェースは，図 3.8 に示すようにタッチペンによる入力可能なディスプレイと，ディスプレイの反対側を写すように取り付けられたカメラから成り，ロボットに取り付けた二次元のマーカをカメラで読み

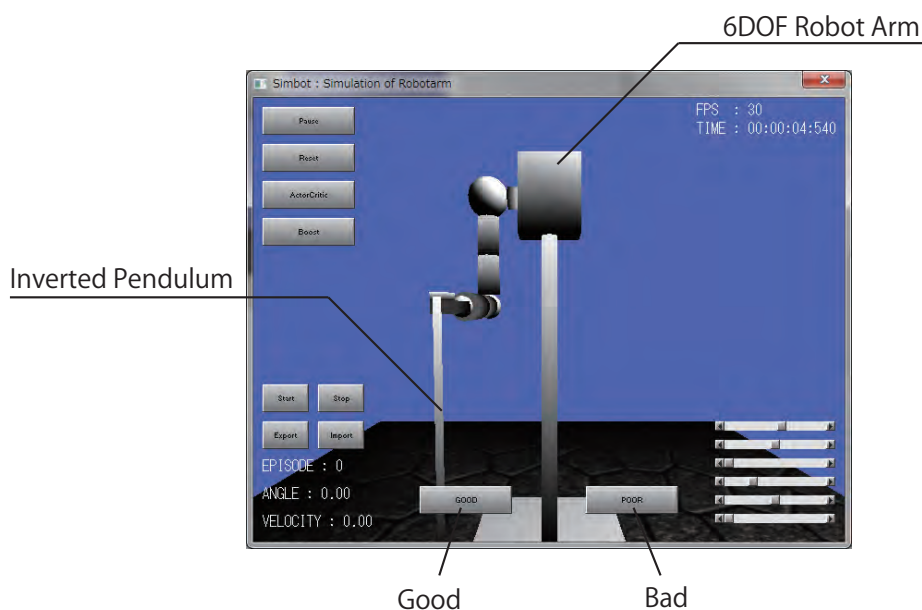


図 3.7: Simulation model

取ることで教示対象を識別する。教示者はタッチペンを用いて「○, ×」などの簡単な記号を画面内のロボットに上書きし、SVMによって手書き文字を認識後無線通信によって二種類の教示が実時間で実ロボットにフィードバックされる仕組みとなっており、以降の実ロボットを用いた検証実験での教示はこのインタフェースを用いて行った。

3.5 評価関数と学習効率

3.5.1 評価関数の分類

ここで、一般的な行動学習手法において評価関数の違いが学習効率に与える影響について検証する。評価関数は一般的に、エージェントによる報酬の獲得機会という観点から、以下の3つの種類に分類できると考えられる。

1. 全状態空間上で報酬が得られる (r_1)

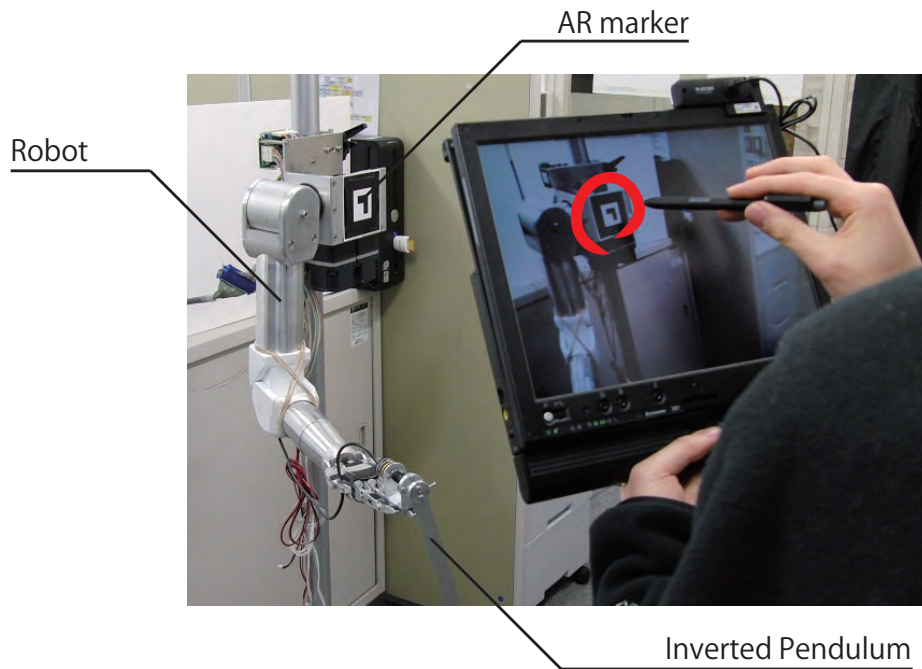


図 3.8: The input interface for teaching real robots.

2. 報酬を得られる領域と、得られない領域がある (r_2)
3. 目的状態でのみ報酬が得られる (r_3)

全状態空間上で評価関数が定義されている場合、エージェントは学習開始時点から方向性のある学習を行うことが出来るが、エージェントが取り得る全ての状態に対して評価関数を予め定義することは通常困難である。一方、エージェントが目標とする状態に到達したときのみ報酬を与えるような評価関数は設計が容易であるが、エージェントの学習効率は著しく下がることが予想される。そのため、通常ではタスクの達成に重要な状態や、目標状態に近い状態などに設計者が経験的に報酬を設定するのが一般的である。

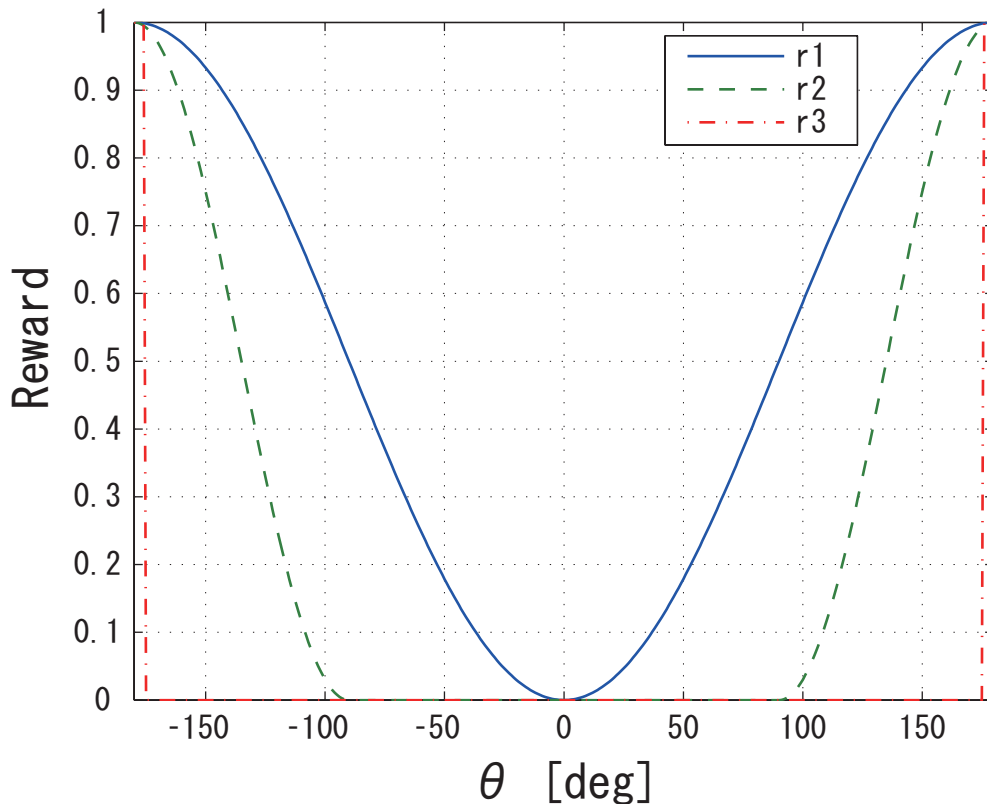


図 3.9: Reward functions for the validation experiment.

3.5.2 検証実験

次に、それぞれの評価関数を用いた場合での学習効率を比較するため、図 3.9 に示す以下の式で定義される評価関数を用いて倒立振子の振り上げ・安定化動作の学習を行った。その結果を図 3.10 に示す。グラフの横軸がエピソード数、縦軸が 1 エピソードの間、振り子の角度が目標角度から $\pm 5^\circ$ 以内にあるときの時間を積算した値を表す。このグラフから、同じ性能の学習エージェントを用いたにもかかわらず、評価関数の特性によって学習効率に大きな差が見られる。特に r_3 においては 100 エピソード後もタスクが達成されなかったため、非常に困難な課題であると言える。その理由としては、この実験ではエージェントが観測可能な状態変数は振り子の角度 θ と角速度 ω としており、この状態空間を広く探索するた

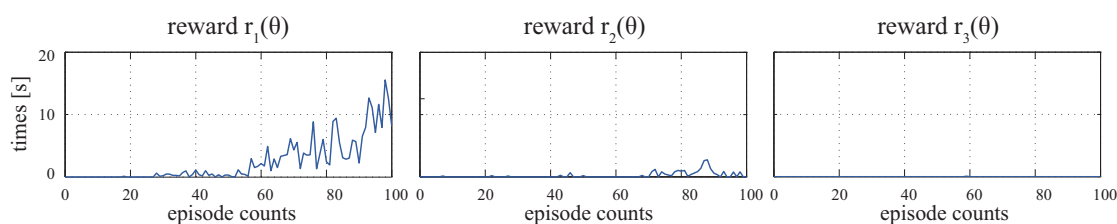


図 3.10: Results of the validation experiment.

めには、振り子の慣性を利用しながら往復運動を繰り返すような行動を徐々に学習しながら探索範囲を広げていくことが重要となるためだと考えられる。

一般に強化学習による行動学習は、学習開始直後はランダムな行動政策によって状態空間の探索を行い、報酬に辿り着いた時点でその情報を手掛かりとして学習を進めていく。しかし、 r_2, r_3 の評価関数では、初期状態周辺において報酬が得られないため上記のような行動が学習できず、結果として学習効率に大きな差が出たと考えられる。このことから、一般的な行動学習において評価関数を設計するときには、 r_1 のように適切な評価関数を設計することが望ましく、これは通常、設計者が試行錯誤によって行うが、目的とするタスクが複雑になるほど望ましい報酬の設計は困難となる。このため、適用する課題を実際にエージェントに実行させながら、主観的なフィードバックを与えることで評価関数を段階的に更新する提案手法は、教示者の負担を減らしつつ機械系の学習を支援するアプローチは妥当であると言える。

3.6 離散系における行動学習支援

3.6.1 離散系の強化学習への実装

はじめに、一般的に学習が比較的容易とされる状態・行動空間が離散的に定義された強化学習に対し、提案手法の実装を行いその有効性を実験によって検証する。ここでは、教示者による教示が与えられた際に以下のように報酬関数の更新を行う。

時刻 t において教示者から評価が与えられたとき、人の教示特性の計測実験から得られた時定数 T_1, T_2 に基づいて、要素数 n の教示の対象となる状態ベクトルの集合

$$X = \{\mathbf{x}_{t-T_2}, \mathbf{x}_{t-(T_2-1)}, \dots, \mathbf{x}_{t-T_1}\} \quad (3.1)$$

を得る。この状態ベクトルの集合に対し、EM アルゴリズムによる混合正規分布モデルを用いた確率の最尤推定を行い確率分布 $P(X|\theta)$ を得る。また、教示の整合性を判断するために過去数回分の教示により作成された $P(X|\theta)$ の積集合を求めることで、類似した状況での教示の一貫性を判断し誤学習を防止する。

$$r(\mathbf{x}) = \min_{i \in C} e^{-\frac{\tau}{T}} P_i(\mathbf{x}|\theta) \quad (3.2)$$

ここで、 $e^{-\frac{\tau}{T}}$ は時間経過による減衰係数であり、 τ は教示が与えられてからの経過時間、 T は減衰時定数、 θ は EM アルゴリズムによって推定した混合正規分布のパラメータ集合を表す。

次に、得られた報酬を用いてエージェントは以下のように状態価値の更新を行う。

$$\delta = r + r_{init} + \gamma V(s_i) - V(s_{i-1}) \quad (3.3)$$

$$V(s_i) \leftarrow V(s_i) + \alpha \delta \quad (3.4)$$

ここで r_{init} は、エージェントに予め組み込まれる報酬関数で、提案手法によって更新される報酬関数と区別して初期報酬関数と呼ぶ。また、 s_i は離散化された状態、 $V(s_i)$ は状態 s_i の状態価値をそれぞれ表す。

3.6.2 検証実験

実験は振り子が鉛直下向きの状態から始め、1回30秒間のエピソード中に振り子を倒立状態から $\pm 5^\circ$ の範囲まで振り上げたら成功とする。状態空間の定義域はそれぞれ $-180[deg] \leq \theta \leq 180[deg]$, $-600[deg/sec] \leq \omega \leq 600[deg/sec]$ とし、シミュレーション時は 36×20 、実機時は 18×10 に等間隔に分割することで状態

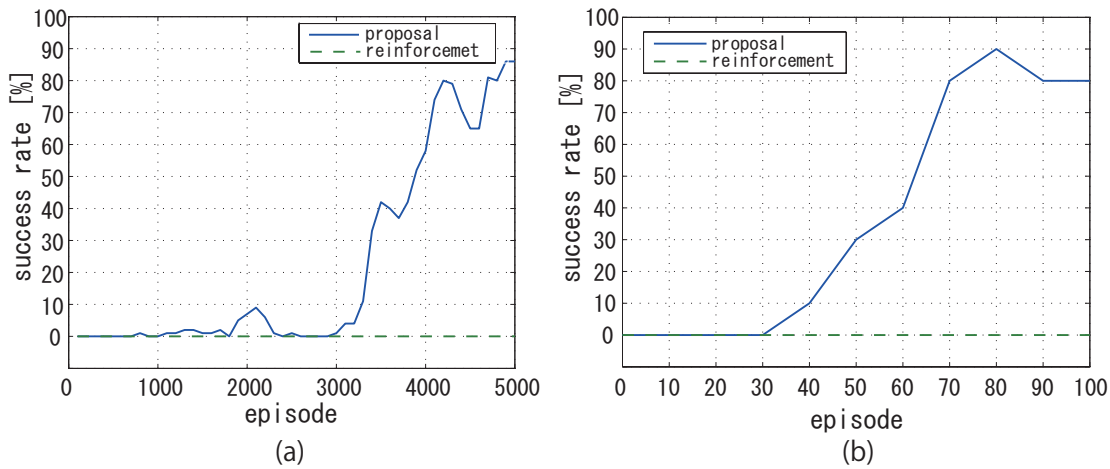


図 3.11: Results of the comparison experiment in a discrete environment.

空間の離散化を行った。

シミュレーション及び実機による実験結果を図 3.11 に示す。各グラフの横軸はエピソード数を表し、縦軸は図 3.11(a) では 100 エピソードごと、図 3.11(b) では 10 エピソードごとの成功回数をエピソード数に対して百分率で表し、成功率として評価している。本実験ではシミュレーションと実機実験で試行回数のスケールが大きく異なるため、倒立維持時間の積算による評価ではなく一定回数ごとの成功率によって提案手法の有効性の検証を行った。この結果より、強化学習のみでは 1 度もタスクを成功しなかったのに対し、提案手法を用いた場合ではシミュレーションと実機の両方で学習が成功していることがわかる。

また、提案手法の妥当性を示すために、報酬関数の時間的な推移について検証を行った。図 3.12 の各グラフは、提案手法によるシミュレーション実験で、それぞれ 100, 500, 1000, 3000 エピソード経過した時点でのエージェントの状態空間における報酬関数を等高線で表している。この図から、学習が進むにつれて高い報酬を与える状態が初期状態の周辺から目標状態へと徐々に遷移している過程が見取れる。これは、エージェントの進捗状況に合わせて段階的に報酬を与えていった結果であり、このことから提案手法による評価関数の動的更新のアプローチが、

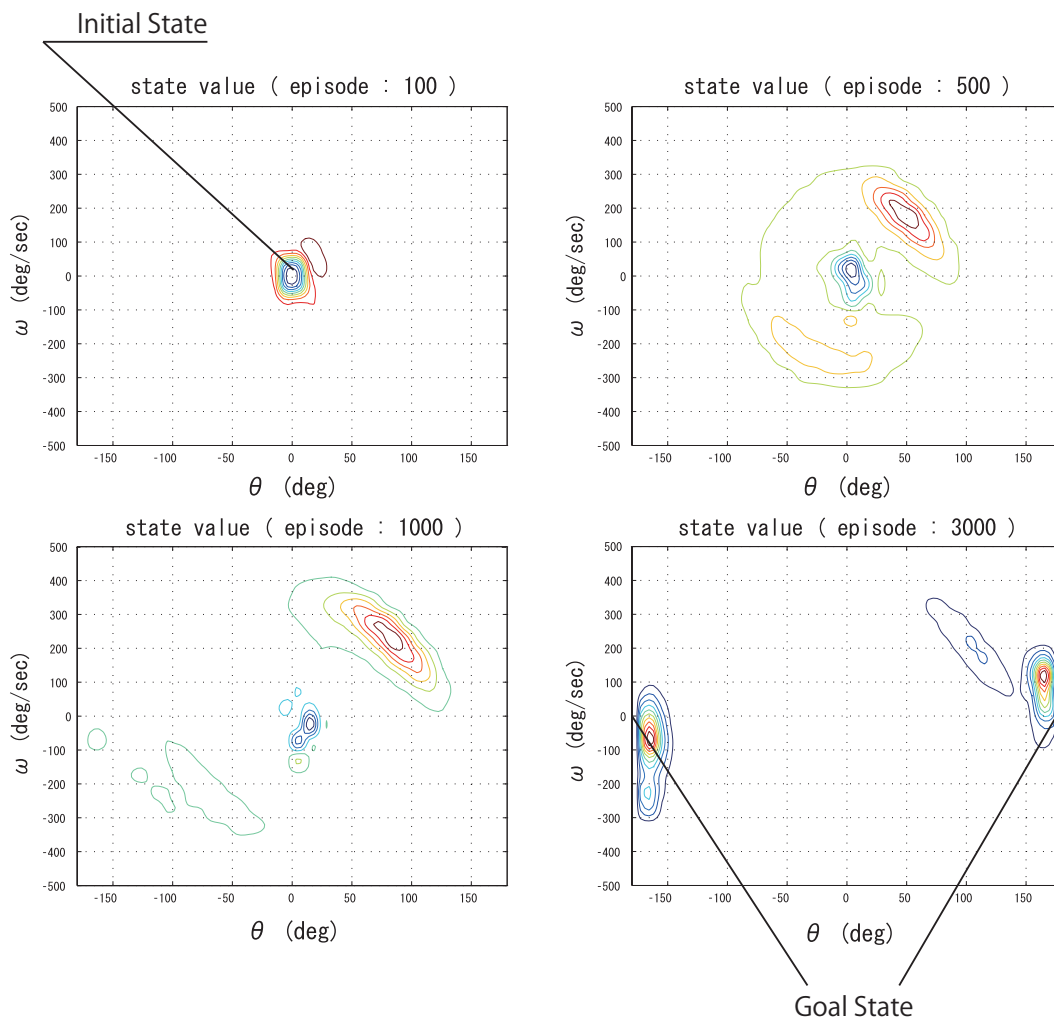


図 3.12: The transition of the reward function.

人がエージェントの学習を支援するための方法として妥当なものであるといえる。

3.7 連続系における行動学習支援

3.7.1 連続系の強化学習への実装

次に状態・行動空間が連続的に定義された強化学習への提案手法の実装について説明する。連続な状態空間上での報酬関数の表現として、以下のようにRBF(Radial

Basis Function) ネットワークによる近似を用いる.

$$r(\mathbf{x}_t) = r_{init}(\mathbf{x}_t) + \sum_k w_k b_k(\mathbf{x}_t) \quad (3.5)$$

$b_k(\mathbf{x})$ は k 番目のユニットの出力, w_k は k 番目のユニットの重みを表す. r_{init} は初期報酬関数を示す. また基底関数にはガウス関数

$$b_k(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \boldsymbol{\mu}_k\|}{\rho\sigma_k^2}\right) \quad (3.6)$$

を用いる.

教示者から「良い」との評価が与えられた場合, X の各状態ベクトルは, タスク達成のために価値の高い状態であると考えられる. そのため, この行動をエージェントに自律的に学習させるため, X の周辺で報酬が得られるようにすることが妥当であると考え, 以下のように報酬関数の重み w を更新する.

$$\delta_k = \lambda^{n-k} R + \gamma r(\mathbf{x}_k) - r(\mathbf{x}_{k-1}) \quad (3.7)$$

$$w_i \leftarrow w_i + \alpha \delta_k b_i(\mathbf{x}_k) \quad (3.8)$$

ただし, R は提案手法による即時的な報酬を意味し, λ ($0 < \lambda \leq 1$) は減衰係数とする. 上記の処理を $k = n$ から $k = 1$ まで繰り返すことによって, \mathbf{x}_n から \mathbf{x}_1 まで緩やかに減衰する勾配が, 報酬関数上に形成される.

一方, 「悪い」との評価が与えられた場合は, X の各状態ベクトルがタスク達成のために価値が高くない状態と言える. これより, X の各要素に対して状態と基底関数の中心との距離が一定値以下なら, その基底の重みを 0 にするという操作を繰り返す.

$$\text{if } \|\mathbf{x}_k - \boldsymbol{\mu}_i\| \leq d \quad (3.9)$$

$$w_i = 0 \quad (3.10)$$

この処理を報酬関数, 状態価値関数, 行動出力関数に対してそれぞれ行うことで

対象の状態周辺の報酬と状態価値を0に戻し、ランダムな行動出力を行うようにする。これは、エージェントの学習が局所解に陥ってしまったときに効果的な方法であり、教示者が誤った教示が行れたと判断した時に素早く修正を行うための方法でもある。

3.7.2 検証実験

ここではエージェントの学習手法として、RBF ネットワークによって状態価値と行動出力の近似を行う Actor-Critic 法を用いる [61][62]。時刻 t における状態ベクトル $\mathbf{x}(t) = \mathbf{x}_t$ に対して、Critic による状態価値関数 $V(\mathbf{x}_t)$ 、Actor による行動出力関数 $u(\mathbf{x}_t)$ を RBF ネットワークの出力として次式のように求める。

$$V(\mathbf{x}_t) = \sum_k z_k b_k(\mathbf{x}_t) \quad (3.11)$$

$$u(\mathbf{x}_t) = \sum_k v_k b_k(\mathbf{x}_t) + n_t \quad (3.12)$$

ただし、 z_k, v_k は各ユニットの重み、 n_t は探索のためのノイズを表す。学習は以下の手順を繰り返すことによって行う。

- i) 現在の状態 \mathbf{x}_t を観測し、Critic により状態価値 $V(\mathbf{x}_t)$ を計算する。
- ii) TD 誤差 $\delta = r(\mathbf{x}_t) + \gamma V(\mathbf{x}_t) - V(\mathbf{x}_{t-1})$ を求める。
- iii) Actor により行動 $u(\mathbf{x}_t)$ を実行するとともに、RBF ネットワークの重みの更新を以下の通り行う。

$$z_k \leftarrow z_k + \alpha \delta b_k(\mathbf{x}) \quad (3.13)$$

$$v_k \leftarrow v_k + \beta \delta u(\mathbf{x}) b_k(\mathbf{x}) \quad (3.14)$$

ただし、 α, β は学習係数である。

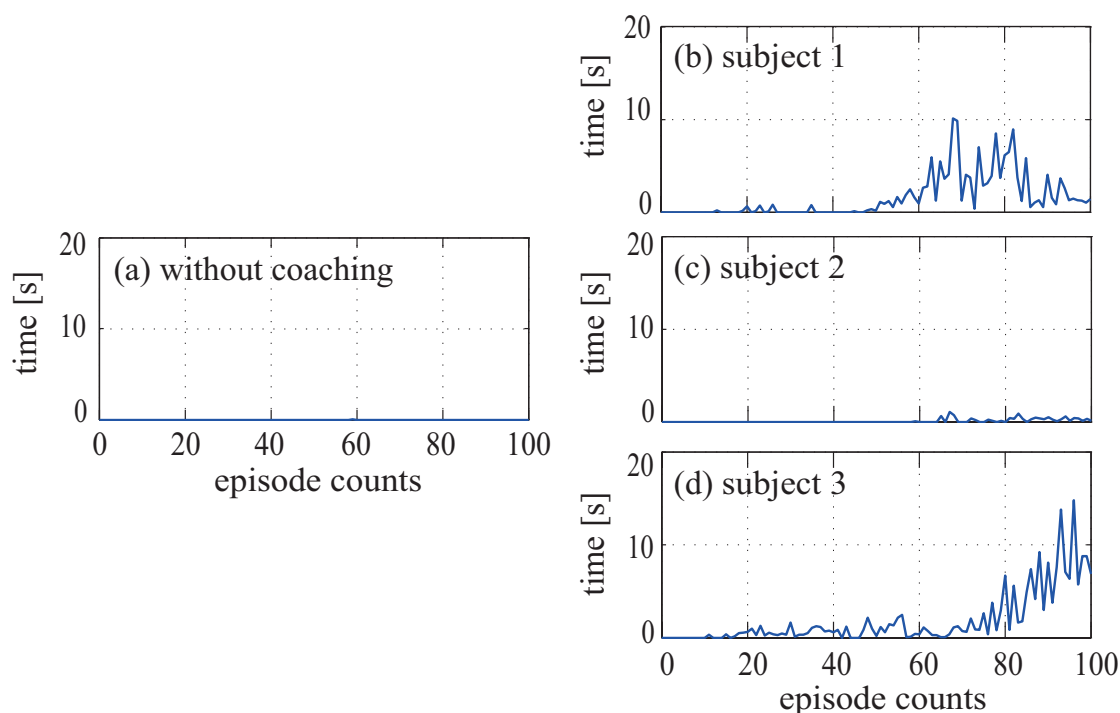


図 3.13: Results of the simulation experiment.

ここでは、振り子が鉛直下向きの状態から始めて 1 回 30 秒の試行中に倒立状態から $\pm 5^\circ$ の範囲にあった合計時間を評価する。実験は提案手法による行動学習支援と通常の強化学習との比較実験とし、図 3.7 に示すシミュレーション環境を用いて実験を行った。

図 3.13 に、通常の強化学習での結果と、3 名の被験者が支援を行った場合の結果を示す。各グラフの横軸がエピソード数、縦軸が倒立状態の時間を表している。まず、提案手法による支援の無い場合では、全試行を通して振り子を目標角度まで振り上げることが出来なかったのに対し、提案手法では 3 人中 2 人の被験者で学習結果に大幅な改善が見られた。一方、図中右下のグラフで示す被験者では学習効率に目立った改善は見られなかった。その理由としては、振り子を振り上げる動作を獲得させるために「勢いよく振り上げる」動作に対して「良い」評価を集中的に与えた結果、振り子を何度も回転させる動作を繰り返す過学習によって、倒立の維持が困難となっていたことが原因であると考えられる。このような場合

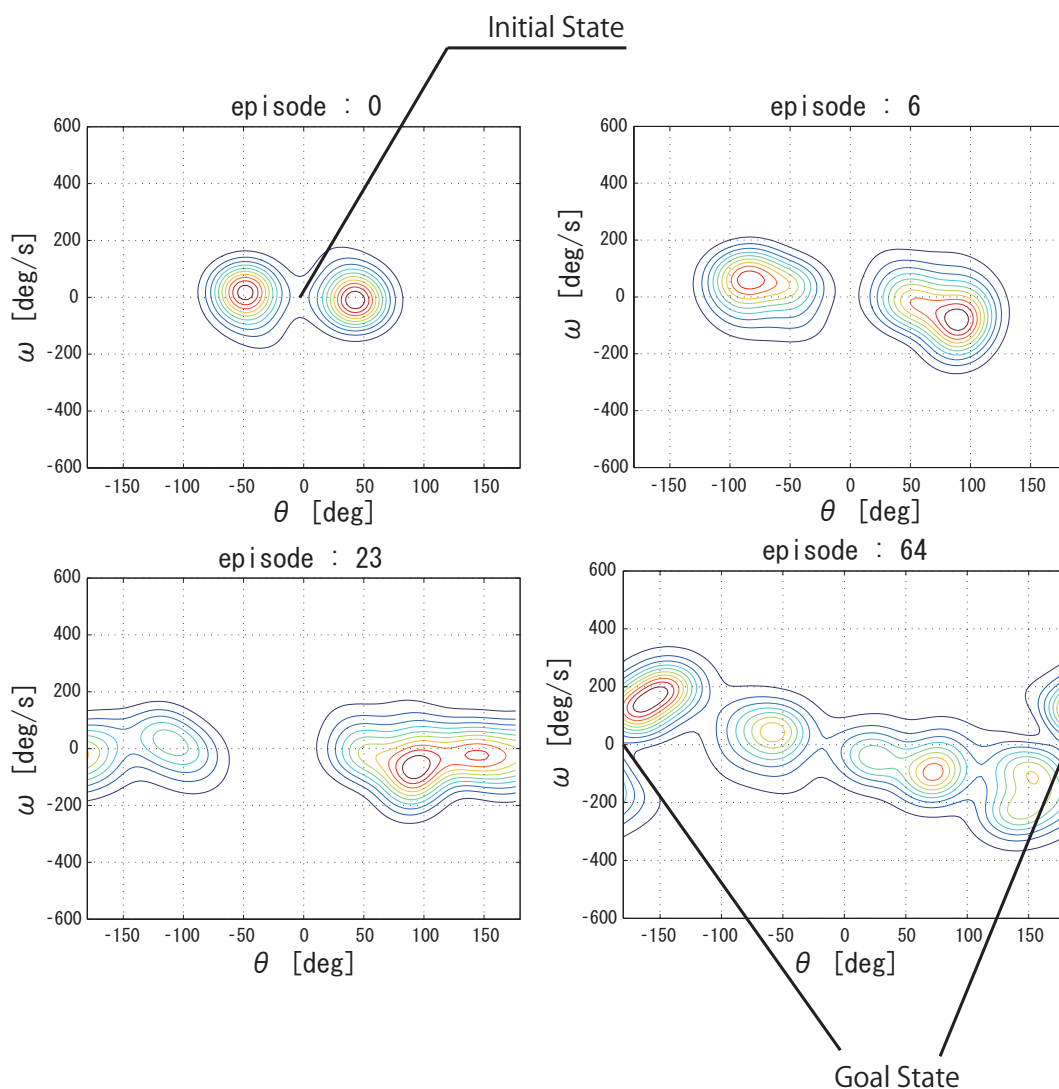


図 3.14: The transition of the reward function in the simulation experiment.

を想定して、提案手法では教示者の判断で負の教示を与えることで以前に与えられた教示による学習効果を低減させることが可能であり、実際に本実験の際にもそのような教示者による教示行動が見られた。このように提案手法では、教示者によって報酬関数を動的に更新可能なため、学習の開始時に報酬関数を固定する従来の手法と比較して、エージェントにより柔軟な学習を行わせることが可能であると考えられる。また、提案手法によって作成された報酬関数の時間的な推移

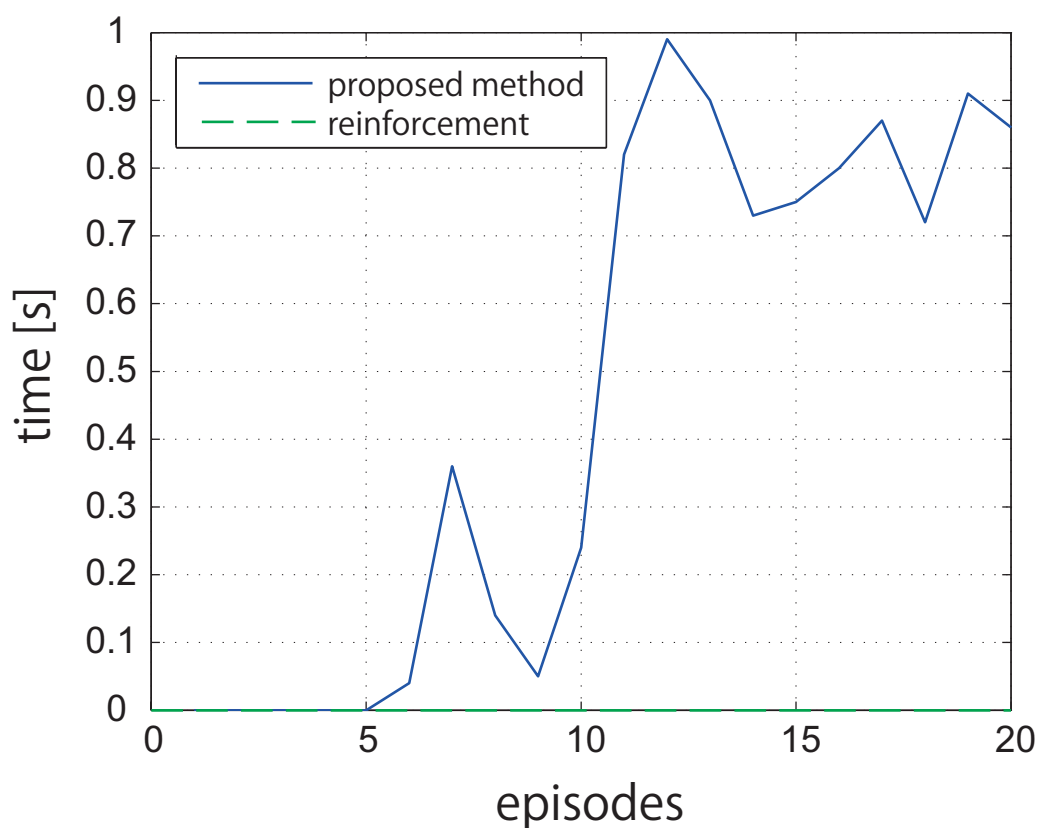


図 3.15: Results of the experiment with a real robot.

についても検証を行った結果，図 3.14 に示すように初期状態から目標状態まで連続的に報酬関数が作成されていることを確認した。

また，実機を用いて同様の実験を行った結果を図 3.15 に示す．この図から分かるように，実機を用いた場合でも提案手法によって強化学習エージェントの行動学習効率を大きく向上させることが出来た．このことから，教示者による主観的な評価に基づく行動学習支援のアプローチ，および報酬関数の動的生成手法の有効性が証明出来たといえる．

3.8 第4章まとめ

本章では，人間機械系の情動的相互作用に基づく行動学習支援の事例研究として，人がロボットの行動学習を支援するための手法を提案した．これは，人が実時間で主観的な評価を与えてロボットの評価関数を動的に更新することで，評価関数の設計に関する既存の行動学習手法の問題を解決するための試みである．ここでは，実時間で与えられる簡便な二種類の教示に基づいて評価関数を適切に更新するために，随伴性・一貫性という人の教示特性に着目して教示意図を推定する手法を考案した．それによって通常の強化学習では学習が困難な問題についても，提案手法によって評価関数を動的に更新することで効率的に学習が可能になることを，シミュレーション実験および実機を用いた実験によって明らかにした．

第4章 社会的相互作用に基づく行動学習支援

4.1 事例研究の位置づけと概要

最後に、本章では人間機械系の社会的相互作用に基づく行動学習支援の事例研究として取り組んだ、自閉症スペクトラム障害を持つ児童のためのロボット介在活動について述べる。

4.2 背景

4.2.1 自閉症スペクトラム障害

自閉症スペクトラム障害 (Autism Spectrum Disorders, ASD) とは、相互的な社会関係とコミュニケーションのパターンにおける質的障害、および限局した常同的で反復的な関心と活動の幅によって特徴づけられる一群の先天的障害と定義されており [63]、診断手法の改善や社会的な認知度の向上により近年世界的な罹患率の上昇が見られる [64][65]。このため、ASD を持つ人々の日常生活における社会的なコミュニケーションが必要とされる場での QOL の向上が重要な課題となっており、また早期の診断と介入がより効果的であることも実証されてきていることから [66]、ASD を持つ児童 (ASD 児) を対象とした社会性の向上やコミュニケーションスキル獲得のための数多くの療育手法が提案されている。

4.2.2 療育(支援)ロボットの意義

一般に ASD 児の特徴として、人や動物など動作の予測が困難なものには関心を示さない、または忌避するが、単純な構造を持つ機械やおもちゃなどの人工物に対して強い興味を示すことが知られている [67] [68]。そのため、ロボットを始めとする機械システムを ASD 児の療育に積極的に取り入れる試みが進んでおり、人に代わって療育を行う、または人の療育者を支援するための様々なロボットが開発されている [69]。特に、人と似通った身体構造を持ち、物体としての親しみ易さと人間らしい存在感の両方を兼ね備えたヒューマノイドロボットは、ASD 児の興味を引きつけるための単なる道具ではなく、人工物と生物の中間的存在として ASD 児のコミュニケーションの幅を広げるための重要な役割を担う可能性が期待されている。

4.3 関連研究と問題点

4.3.1 従来の療育手法

ASD 児の療育手法として最も有名な手法として、Skinner らの提案した応用行動分析 (Applied Behavior Analysis, ABA) が挙げられる [29]。これは、第2章でも述べた自発的な行動に随伴する報酬または罰による行動の強化・弱化的メカニズムを、ASD 児の行動変容を目的とした訓練に応用したものである [70]。ABA は人や動物が生得的に持つ学習メカニズムを利用するため目に見える行動変容の効果が強く、また療育者による直観的な理解も容易であり特別な設備も必要としないことから数多くの療育現場で実践的に用いられている。ここで、ABA によって変容可能な行動は本質的に被療育者が自発的に行うことができる行動のみであり、そのため多くの場合において ABA は ASD 児がしばしば行う社会的に受け入れられない行動 (以下、ネガティブ行動) の低減のために用いられているが、ABA によって単にネガティブ行動の頻度を減らすだけでは不十分であり、社会的に好意的に受け取られる行動 (以下、ポジティブ行動) の頻度を同時に高めることが重要性と



図 4.1: Animal-Assisted Activities

いわれている。そこで、動物介在活動 [71] や SCERTS モデル [72] など、ASD 児の情動と行動の関係に着目し、ASD 児が快情動を感じることで他者との触れ合いやコミュニケーションに安心や楽しさといったポジティブな主観的価値を見出し、それが他者との社会的交流のモチベーションにつながることを目指す児童主体の療育アプローチも提案されている。

4.3.2 動物介在活動

動物介在活動 (Animal-Assisted Activity, AAA) は、動物とのふれあいを通じて ASD 児の快情動を誘発することを目的とした療育の形態であり [71] [73]、他者に対するポジティブ行動の促進や社会性が向上した例が数多く報告されている [74] [75]。しかしながら、AAA では生きた動物を用いるためアレルギーやトラウマ体験を持つ児童の参加は困難であり、またセラピーアニマルとしての育成や管理に時間と費用が掛かることから実施出来る環境は限られる。さらに、動物の行動を療育者がコントロールすることが出来ないため、療育者の望む時点で適切な介入や繰り返しは困難であることから、介入の効果が動物の特性や児童との相性に極端に左右され易いなどの問題点が指摘されている。

4.3.3 療育(支援)ロボットに関する先行研究

療育(支援)ロボットに関する大多数の先行研究では、達成したい社会的行動(例えば、挨拶やアイコンタクト等)を標的行動として設定し、児童とロボットとのインタラクションの中でシンプルで分かり易い教示を用いて ASD 児による主体的な標的行動の実行を促すとともに、光や音、動作等による報酬をロボットが提示することで標的行動を強化するといった、ABA に立脚した従来の療育プログラムをロボットを用いて実施するアプローチが主流となっている [76][77]。しかしながら、このアプローチでは ASD 児は基本的にロボットと一対一の状況で、目的とする行動をそれが本来使用されるべき社会的な文脈から切り離して繰り返し学習するため、獲得した社会的スキルのその他の状況における汎用性という点に疑問が残る。

また、これらの先行研究ではロボットと児童のインタラクションのシナリオやパターンは予め決められており、それに基づいて事前にプログラムされた複数の行動セットの中から、療育者が状況に合わせてロボットの適切な行動を選択的に実行する、Wizard of OZ(WOZ)方式が一般的である [27]。一方、人による操作を必要としない自律型の療育ロボットに関する研究も行われているが [78][79][80]、療育において重要とされる視線や表情、ジェスチャーなど児童の様々な社会的行動指標を認識し、適切な応答を自律的に返すようなロボットの開発は非常に困難であることから、現状では WOZ 方式を用いることが最も妥当であるといえる。

ASD を持つ児童の自閉性傾向や現有スキルは、自閉症“スペクトラム”と定義されるように様々に異なるため、一般的に療育プログラムは対象児童を良く知る医者やセラピストなど複数の専門家によって個々の児童に合わせて個別に策定される。しかしながら、このような事前に用意した行動セットを用いる WOZ 方式では事前のシナリオを逸脱したやり取りには適切に対応することが出来ず、また療育者によるロボットの行動の即興的な作成や変更も困難であることから、療育的な観点からも問題点が指摘されている [27]。

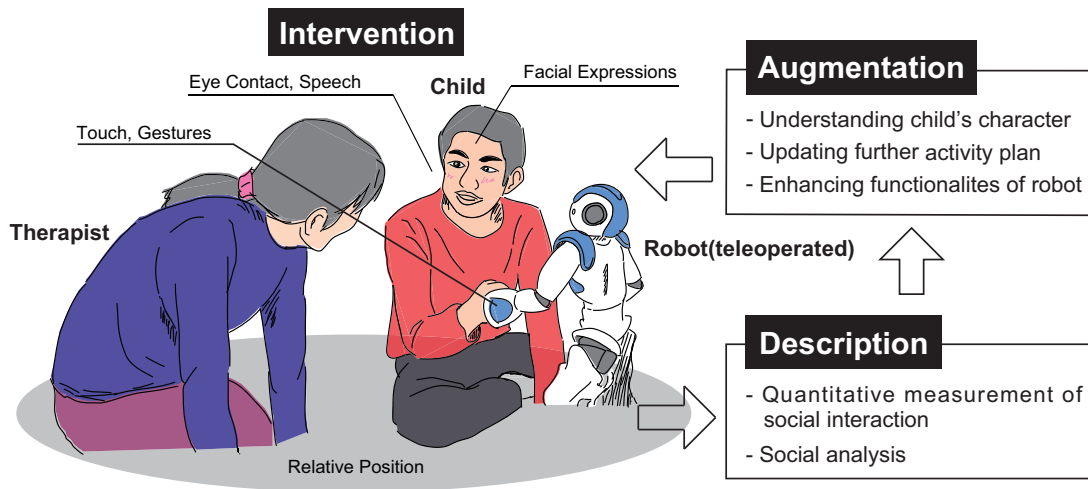


図 4.2: Proposed Robot-Assisted Activity.

4.4 提案手法

4.4.1 個人適応的ロボット介在活動

そこで本研究では動物介在活動に着想を得て、図 4.2 に示すような ASD 児とロボットとの社会的なインタラクションを通じて快情動を誘発することで、ASD 児の自発的な社会的行動の学習を支援するためのロボット介在活動 (Robot-Assisted Activity, RAA) を提案する。ここでは、療育者による実時間での直観的なロボットの操作を可能にする人形型インタフェースを用いた制御手法を提案する。それによって、既存の WOZ 方式による制御では困難な動作の作成や修正を児童の反応を観察しながらその場で行うことが出来るとともに、装着型表情計測でバイスによる快情動の推定と、環境センサを用いたインタラクションの定量的記述と組み合わせることで、個々の児童の特徴や療育の進度に合わせて介入のアプローチやプロトコルを柔軟に拡張・修正する、個人適応的なロボット介在活動のためのフレームワークを構築する。

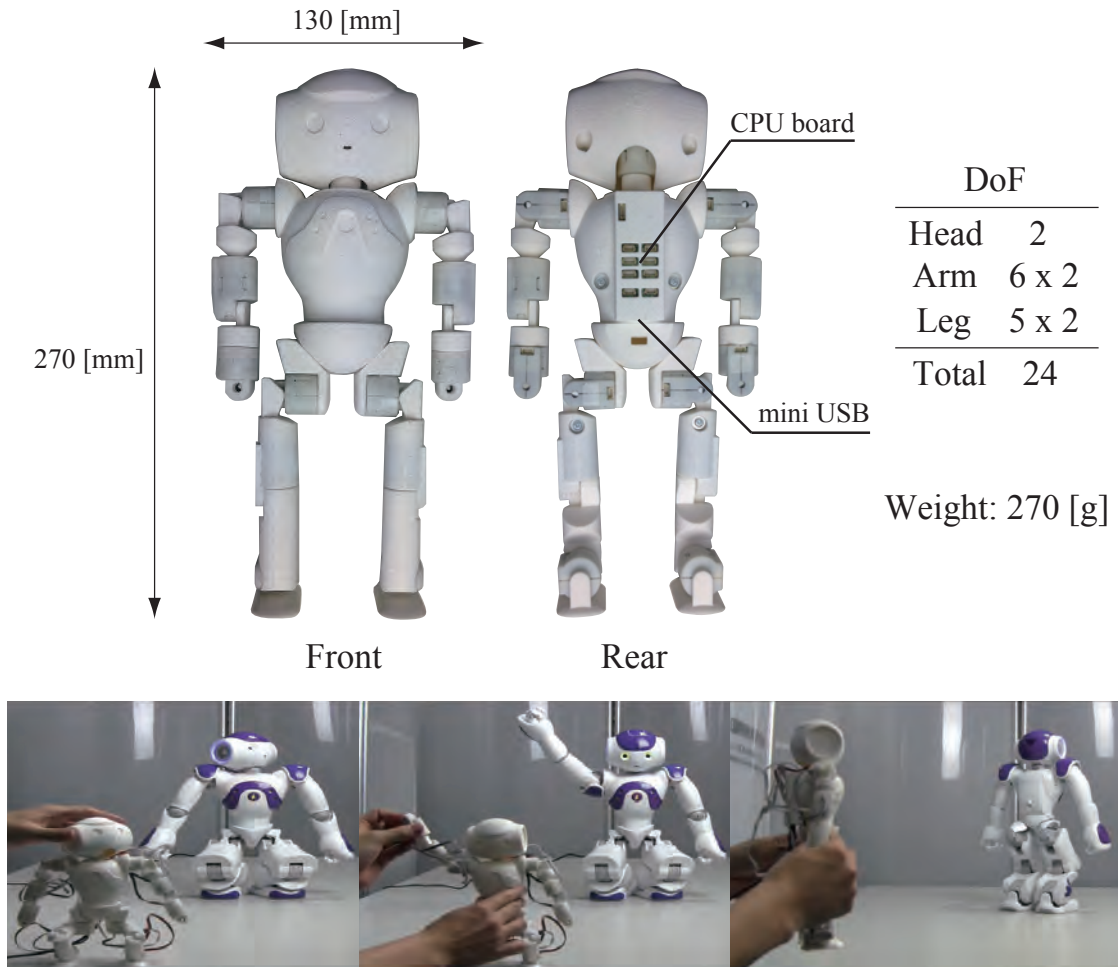


図 4.3: A doll-type robot control interface.

4.4.2 ロボットの実時間遠隔操作手法

人形型インタフェース

ロボットの実時間での遠隔操作手法として，図 4.3 に示すような制御対象のロボットと同様の身体性を有する人形型インタフェースを開発した．今回，操作対象のロボットとして小型のヒューマノイドロボット (NAO, Aldebaran Robotics) を用いることとし，構築する人形型インタフェースの自由度数および関節の軸配置は実際のロボットと等しく成るように設計した．インタフェースの各関節角度は

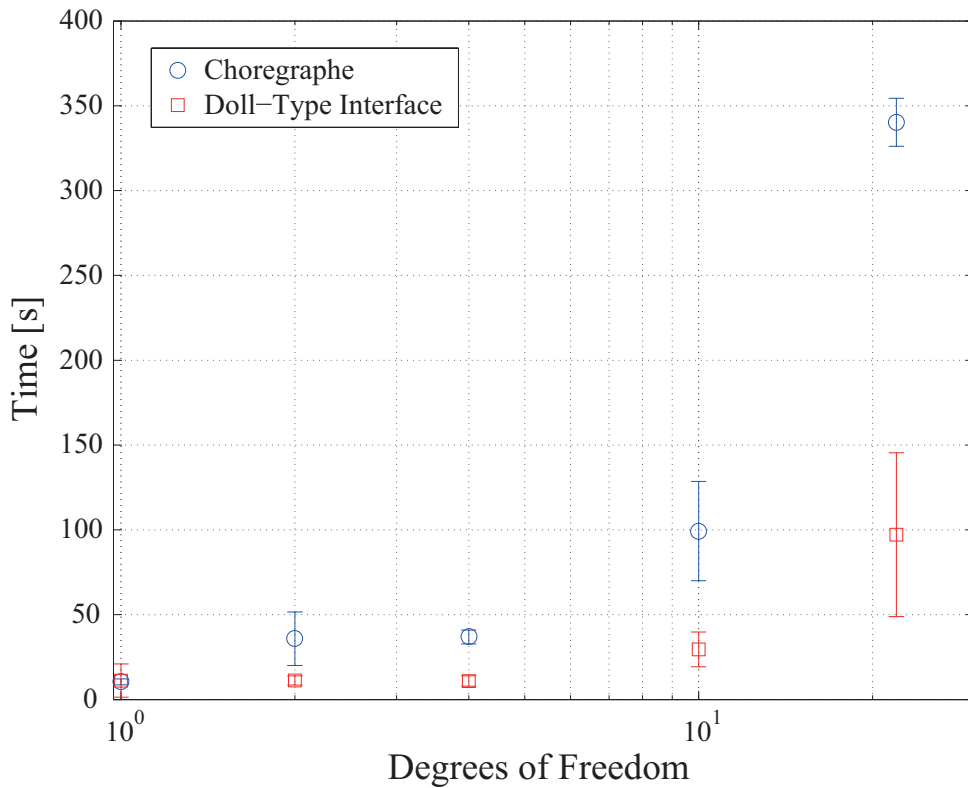


図 4.4: The result of the comparison experiment between the proposed doll-type interface and conventional GUI based interface.

内蔵した角度センサによって約 100 [ms] 周期で取得可能であり、ロボットの関節角はインタフェースの対応する関節を操作することでリアルタイムに制御することができる。また、複数の関節の同時制御にも対応しており、ロボットの非専門家でも複雑な動作を直観的に実時間で入力することができる。これによって、事前の予測が困難な ASD 児の振る舞いにも療育者が即興でロボットを操作することで対応できるとともに、個々の児童の特性や傾向に合わせた動作の修正も任意の時点で行うことができる。

ユーザビリティテスト

本インタフェースを用いたロボットの姿勢入力の有効性を検証するためのユーザビリティテストを行った。ここでは、従来のGUIに基づく制御ソフトウェアと開発したインタフェースとの比較実験とし、複雑度が異なる5種類の姿勢をそれぞれの手法で入力した場合の所要時間を比較する。図4.4に、4人の被験者による結果の平均と標準偏差を示す。この結果から、入力する姿勢が複雑になるほど提案する人形型インタフェースの有効性が明確に現れているのが見て取れ、ヒューマノイドのような複雑多自由度なロボットの姿勢入力方式として妥当であるといえる。

バランス制御手法

しかしながら、インタフェースによって入力された姿勢をそのままロボットに反映させるだけでは、バランスを保つことが出来ず転倒するのは明らかである。そこで、Sugiharaらの考案した重心ヤコビアンに基づくヒューマノイドロボットの動作生成手法 [81][82] を応用したバランス制御手法を開発した。これは、ロボットの関節角速度 $\dot{\theta}$ と重心速度 \dot{p}_G との間の微小変位の関係を表す重心ヤコビアン J_G を用いて、インタフェースによる入力姿勢との誤差を最小にしながらロボットのバランスを保つために必要な重心加速度を生じさせる関節角入力を解析的に求める手法である。

重心ヤコビアンを介した関節角速度と重心速度の関係は次のように表せる。

$$\dot{p}_G = J_G \dot{\theta} \quad (4.1)$$

ロボットのバランスを保つための目標 ZMP(x_Z, y_Z, z_Z) が与えられたとき、目標重心加速度 (${}^{ref}\ddot{x}_G, {}^{ref}\ddot{y}_G, {}^{ref}\ddot{z}_G$) は、現在の重心の三次元空間中の位置を (x_G, y_G, z_G),

ロボットの足裏に掛かる垂直床反力を f_z 、質量を m として次のように得られる。

$${}^{ref}\ddot{x}_G = \omega^2(x_G - x_Z) \quad (4.2)$$

$${}^{ref}\ddot{y}_G = \omega^2(y_G - y_Z) \quad (4.3)$$

$${}^{ref}\ddot{z}_G = \frac{f_z}{m} - g \quad (4.4)$$

また、 ω^2 を以下の通りに定義する。

$$\omega^2 = \frac{f_z}{z_G - z_Z} \quad (4.5)$$

ここで、重心に掛かる加速度を高い精度で直接制御することはハードウェアの性能上困難であることが多いため、実際の制御では以下の式を用いて現在の重心速度と目標重心加速度から求める目標重心速度 ${}^{ref}\dot{p}_G$ について制御を行う。

$${}^{ref}\dot{p}_G = \dot{p}_G + {}^{ref}\ddot{p}_G \Delta t \quad (4.6)$$

得られた目標重心加速度と目標関節角速度 ${}^{ref}\dot{\theta}$ の関係は重心ヤコビアンを用いてが以下のように表せることから、

$${}^{ref}\dot{p}_G = J_G {}^{ref}\dot{\theta}. \quad (4.7)$$

目標関節速度は次に示す二次計画問題を解くことで解析的に求めることが出来る。

$$\frac{1}{2} \left\| ({}^{cmd}\dot{\theta} - {}^{ref}\dot{\theta})^T W ({}^{cmd}\dot{\theta} - {}^{ref}\dot{\theta}) \right\| \rightarrow \min \quad (4.8)$$

$$\text{subject to } p_G = J_G {}^{ref}\dot{\theta} \quad (4.9)$$

W は重み行列を表し、 ${}^{cmd}\dot{\theta}$ はインタフェースから取得した操作者による関節の入力角度と現在の関節角から以下の式を用いて得られる関節入力角速度を表す。

$${}^{cmd}\dot{\theta} = \frac{{}^{cmd}\theta - \theta}{\Delta t}. \quad (4.10)$$

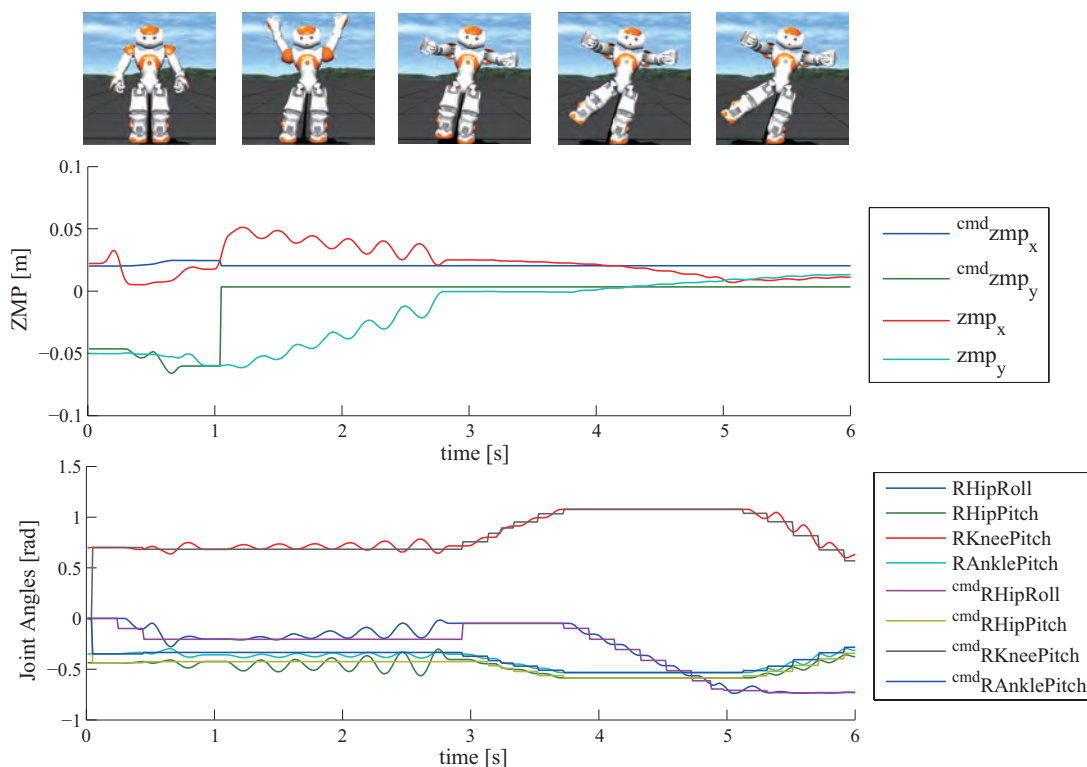


図 4.5: Demonstration of the balancing algorithm.

ここで、実際のロボットに実装する場合には関節の可動範囲や床面との接触などを拘束条件を考慮する必要がある。これらの拘束条件は関節角度とのヤコビアンを用いて以下のように表せることから、

$$p_C = J_C^{ref} \dot{\theta}. \quad (4.11)$$

式 4.9 で示した拘束条件式と合わせて、

$$p_U = \begin{bmatrix} p_G \\ p_C \end{bmatrix} = \begin{bmatrix} J_G \\ J_C \end{bmatrix}^{ref} \dot{\theta} = J_U^{ref} \dot{\theta} \quad (4.12)$$

として二次計画問題を解く。

この手法を用いたロボットのバランス制御の例を図 4.5 に示す。この図から、入

表 4.1: Leg states and the recognized motion

Left leg state	Right leg state	Motion
Extended	Extended	Stand up
Bent	Bent	Sit down
Walking	Walking	Walk forward
Extended	Walking	Turn left
Walking	Extended	Turn right

力姿勢に合わせてシステムが自律的に目標 ZMP を設定しバランスを保つとともに、可能な限り入力関節角と一致するように制御を行っていることが分かる。

操作入力認識

また、操作者による下半身関節に対する時系列の入力パターンを隠れマルコフモデルによって認識し、予めプログラムした動作(起立, 着座, 歩行など)を実行する半自律制御手法も開発した。ここでは、脚部の動作に対して屈曲, 伸展, 歩行の3状態を定義し、左右の脚の状態の組み合わせに応じて表 4.1 に示すような5つの動作を認識することが出来る。モデルを構成するノードの数は10個とし、各ノードの出力は以下のようなロボットの左右の股関節, 膝関節の関節角度, 角速度の4次元ベクトルで表される

$$\boldsymbol{\theta} = [\theta_{Hip}, \theta_{Knee}, \dot{\theta}_{Hip}, \dot{\theta}_{Knee}] \quad (4.13)$$

出力の確率分布 $p(\boldsymbol{\theta}_t)$ は混合ガウス分布 (Gaussian Mixture Model, GMM) を用いて次のように記述する。

$$p(\boldsymbol{\theta}, \xi) = \sum_{i=1}^M N(\mu_i, \sigma_i, \boldsymbol{\theta}) \quad (4.14)$$

M はガウス分布の混合数, μ, σ は各分布の平均と標準偏差, ξ は GMM のパラメータセットをそれぞれ表す。GMM のパラメータの学習には Baum-Welch アルゴリズムを使用し, 10[Hz] でサンプリングされた訓練データを用いて予めパラメータ

の学習を行う。

認識は各時間ステップにおいて過去1秒分の関節角度データを用いて行い、脚部の状態 s はそれぞれの状態に対応するモデルのうち Forward アルゴリズムによって最も高い尤度を示しているモデルの状態が選ばれる。

$$s(t) = \operatorname{argmax} \{p(\boldsymbol{\theta}_t | \xi_{ext}), p(\boldsymbol{\theta}_t | \xi_{bnt}), p(\boldsymbol{\theta}_t | \xi_{wlk})\} \quad (4.15)$$

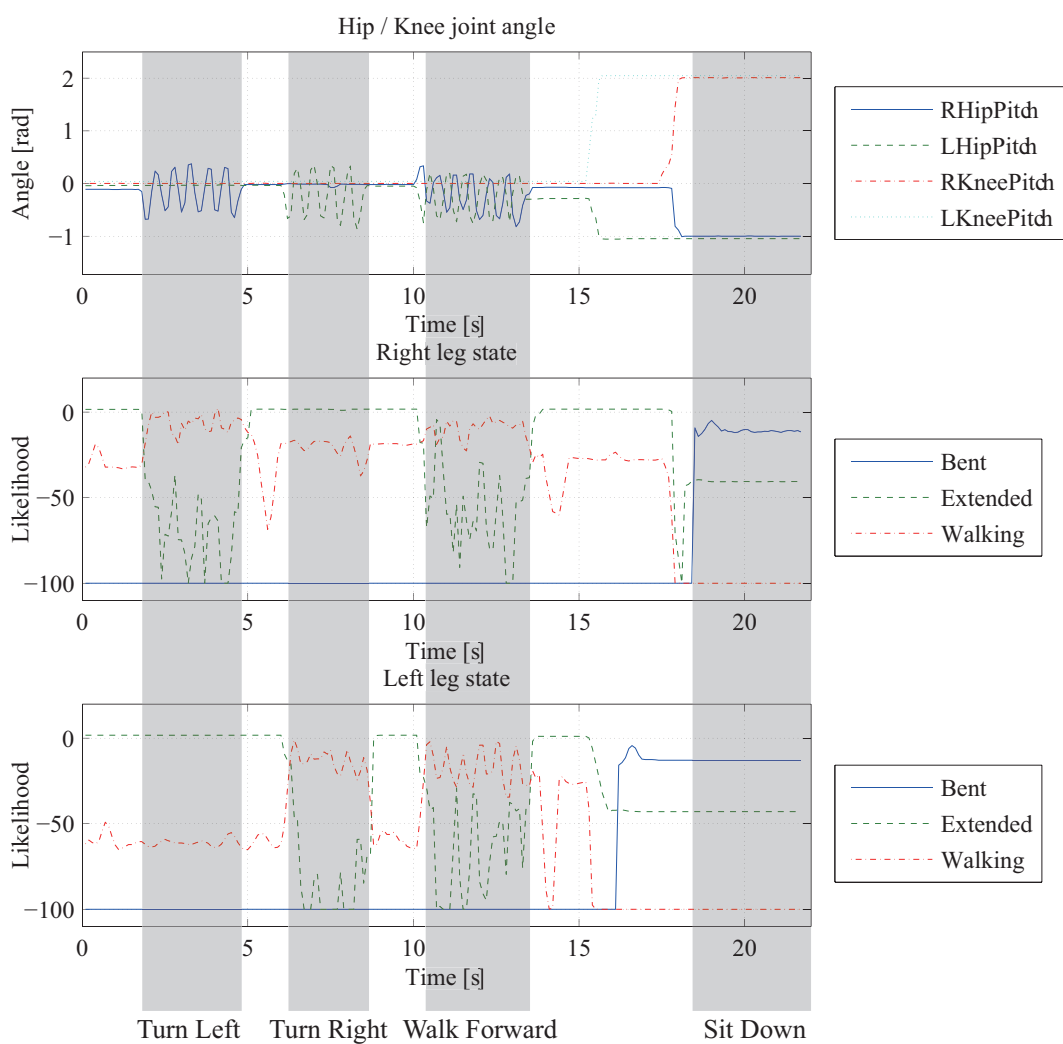


図 4.6: Demonstration of motion pattern recognition.

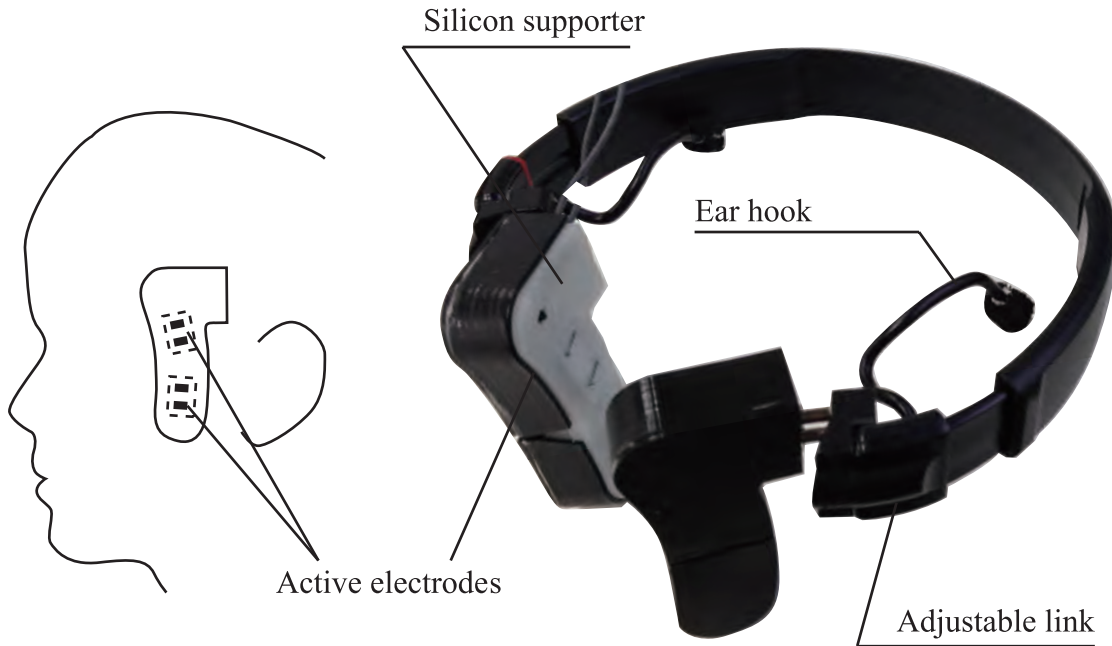


図 4.7: A wearable device for facial expression detection.

図 4.6 に各動作の認識結果の例を示す。この図から、関節の入力パターンに応じて対応する動作の尤度が増減し、適切に認識出来ていることが分かる。

4.4.3 快情動の推定手法

装着型表情計測デバイス

インタラクション中の ASD 児の情動を推定する指標としてロボット介在活動中の児童の表情に着目する。表情は情動と密接に関連する生理的な反応であり、特に笑顔はポジティブな感情の状態を表すことから、療育の効果や対象児童にとっての適性を評価するために計測されることが多い [75]。開発した装着型表情計測デバイスを図 4.7 に示す。このデバイスは表情を形作る表情筋の活動を顔の側面で計測するための乾式電極を備えており、電極の位置を固定するための柔らかいシリコンと、装着者の頭部の大きさに合わせてサイズ調整が可能なヘッドバンドから構成され脱着が容易な仕組みとなっている。本デバイスを用いて左右 2 チャネ

ル, 計4チャンネルの表情筋活動 (Electromyography, EMG) を1[kHz] で無線通信を介してサンプリングが可能である.

表情筋活動に基づく表情識別

サンプリングされた信号から表情識別には Gruebler らが考案した手法を用いる [83]. はじめに帯域通過フィルタを用いて EMG 信号を抽出し, 独立成分分析 (Independent Component Analysis, ICA) を用いて4つの独立成分に変換後, 150[ms] の時間窓での積分値を計算する. 次に, 得られた積分値をニューラルネットワークの入力として表情の識別を行う.

本デバイスを用いることによって, ロボットや療育者とのインタラクションの間の笑顔の時間やその経時的変化, および他の行動指標との関係性を解析することが可能となり, 個々の児童に合わせた適切な介入方法を検討するための重要な有益な情報が得られる.

4.5 評価実験

4.5.1 実験手順

提案するロボット介在活動の検証実験を, 自閉症スペクトラム障害と診断され動物介在活動に一年以上参加している児童6名(男児6名)を対象に, 動物介在活動のセッションが終了した後のポストセッションとして行った(図4.9「保護者の許可を得て記載」). 実験は図4.8に示すようなプレイルーム内で約10~15分程度行い, 被験児とロボットのインタラクションの様子を天井四隅に設置したカメラとプレイルーム内の実験者の持つビデオカメラによって撮影する. プレイルームには観察室が併設されており, ハーフミラーや天井カメラの映像を通してプレイルーム内の様子を観察することが出来る. ロボットの操作には提案した遠隔操作手法を用い, プレイルーム内で被験児を観察するセラピストによるキューに合わせて実験者が操作することで, 被験児の働きかけに対する応答や保護者とのイン

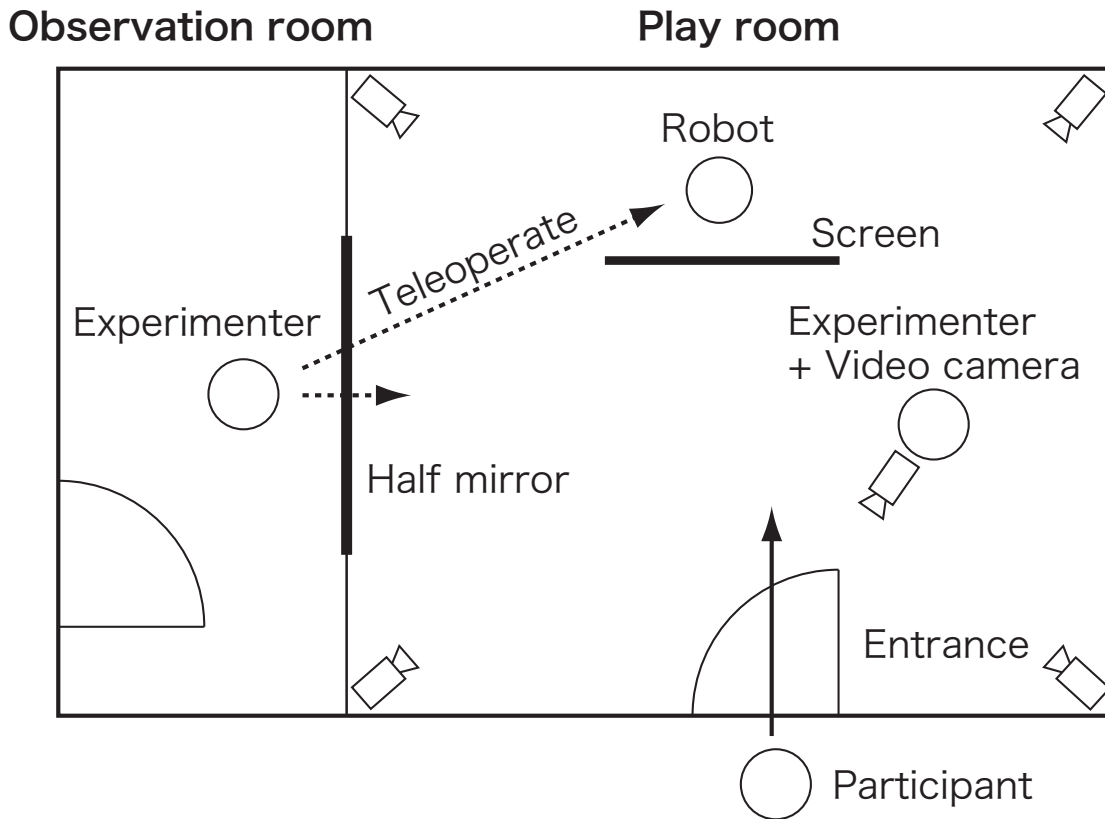


図 4.8: Experiment field.

タラクションを促す.

4.5.2 実験結果

動物介在活動との比較

はじめに、今回の実験に参加した被験児は動物介在活動にも参加していることから、動物介在活動 (AAA) とロボット介在活動 (RAA) のそれぞれの初回のセッションにおける同一被験児の反応を比較した。ここでは、実験者が撮影したビデオデータを用いて療育の専門家が客観的な行動分析を行った結果に基づいて解析を行った。図 4.10 に、ASD 児によるポジティブな社会的行動と専門家が判断した行動の合計時間のセッション時間に対する割合を示す。RAA の初回セッションに

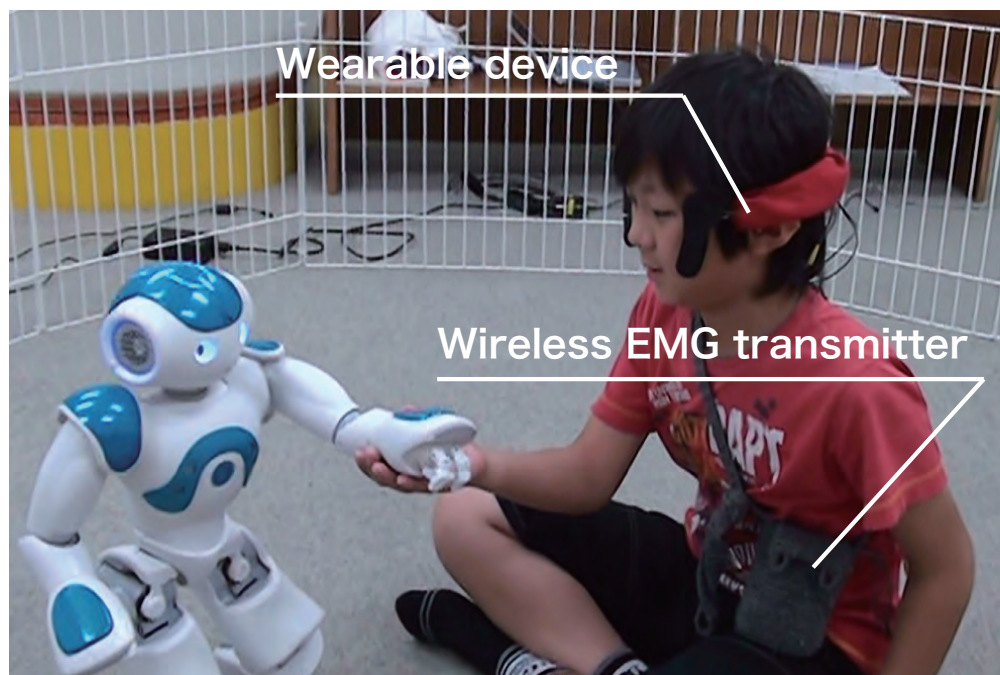


図 4.9: Experiment overview (Normal control participant) .

におけるポジティブな社会的行動の割合が AAA の初回セッションと比べると同程度であり、符号検定を行ったところ二つの場合に有意差は確認出来なかった。

同様に、ASD 児によるネガティブな社会的行動と専門家が判断した行動の合計時間のセッション時間に対する割合を図 4.11 に示す。この図から分かるように、6 名中 5 名の被験者で RAA の初回セッションでのネガティブな社会的行動の割合が AAA に比べて減少しており、AAA でネガティブ行動が観察されなかった残りの 1 名についても変化無しという結果が得られた。この結果について符号検定を行ったところ、危険率 10% で有意に減少していることが分かった。

以上の結果から、RAA の初回のセッションを AAA と比較した場合、ポジティブ行動の割合はそのままに、ネガティブ行動の割合に関して有意に減少していることが分かった。これは、複雑で動きの予測のつかない生物よりも、単純な動作を繰り返す人工物に興味を示すという ASD 児の一般的な傾向に由来するものと思われる。しかしながら、観察されたロボットに対するポジティブ行動の内容は、ボ

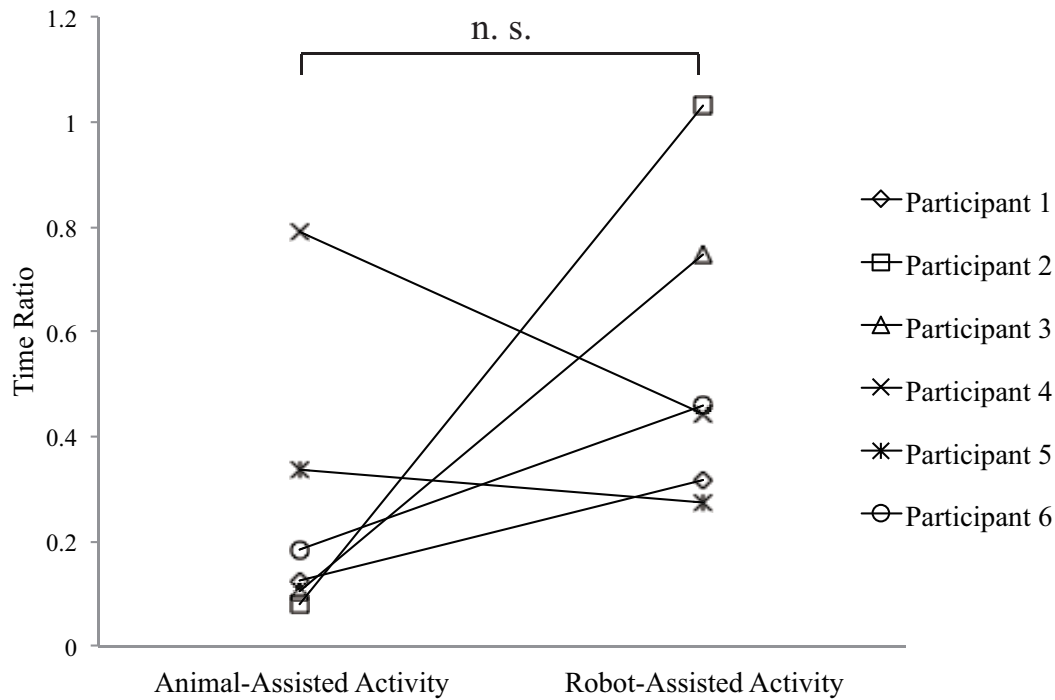


図 4.10: The percentage of time the subject performed positive social behaviors in animal-/robot-assisted activity.

タンを押す、関節の動きを観察するなど、ロボットを単なるオブジェクトと見なして接していた行動は含まれておらず、挨拶や握手、手を振るといった社会的な行動のみ集計している。このことから、ASD 児にとってロボットは生き物ではないが完全な無機物でもない、それらの中間的な存在として位置づけられている可能性が示唆される。

Face-to-Face とポジティブ行動の関係

次に、ロボットの頭部に搭載されたカメラの映像を用いて、ASD 児がロボットの顔を正面から見ている状態 (Face-to-Face) に着目して解析を行った。Face-to-Face や視線を合わせる動作 (アイコンタクト) はコミュニケーションを行う上で非常に重要であり、ASD 児の特徴としてこれらを意図的に回避することが知られている

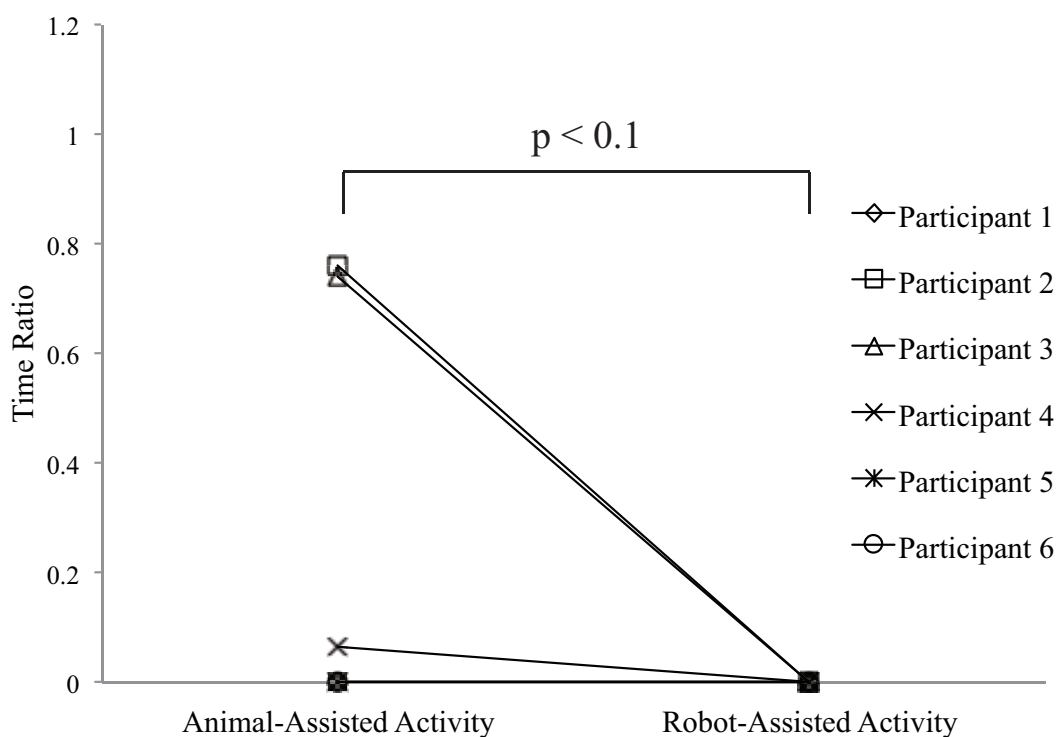


図 4.11: The percentage of time the subject performed negative social behaviors in animal-/robot-assisted activity.

ことから [84], Face-to-Face やアイコンタクトは ASD 児の自閉性傾向や療育の効果を知る指標として考慮されており, これらのコミュニケーションスキルを学習することを目的とした療育ロボットの研究も行われている [28] [78].

ここでは, Face-to-Face と前述した専門家により同定された ASD 児のポジティブ行動との関係について, ポジティブ行動と時間的に近接する ($\pm 1[\text{sec}]$ 以内) Face-to-Face のみを抽出し, その合計時間の全ての Face-to-Face の合計時間に対する割合を図 4.12 に示す. この結果から, 6 名中 5 名の被験者でほぼ全ての Face-to-Face 区間がポジティブ行動と時間的に重なり合っていることが見て取れる. また, 残りの 1 名に関しては一緒に参加した被験者の弟がセッションの間ロボットを占有していたため, Face-to-Face 動作そのものが観察出来なかったことが原因である. この結果より, ASD 児がロボットの顔を見無目的に見ていたのではなく, 挨拶や握手などの

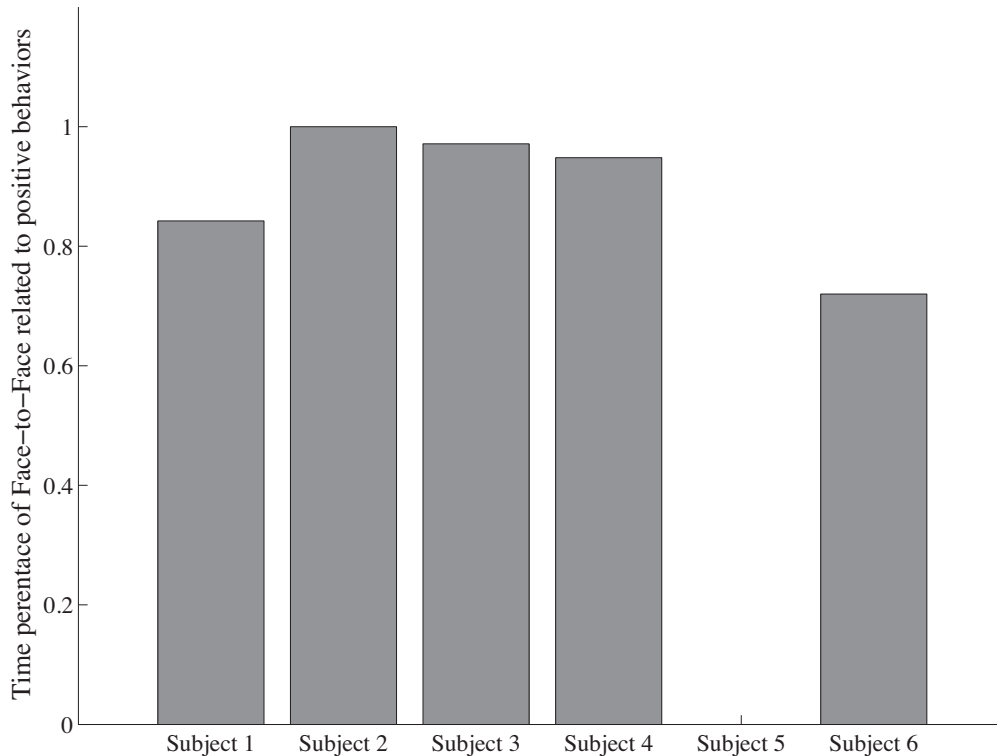


図 4.12: Time percentage of Face-to-Face behaviors related to positive social behaviors.

行動の起点やそれによる相手の反応を伺うという明確な意図を持って Face-to-Face 動作を行っていた可能性が示唆されることから、本研究の目指す自然なインタラクションの文脈を通じたコミュニケーションスキルの自発的な学習の実現可能性も確認出来たといえる。

4.6 第4章まとめ

本章では、人間機械系の社会的相互作用に基づく行動学習支援の事例研究として、自閉症スペクトラム障がいを持つ児童の社会的行動の学習を支援するロボット介在活動について研究を行った。これは、ASD 児のポジティブ、ネガティブ行動の起因として情動に着目し、動物との触れ合いを通じて快情動を誘発することで他者に対する社会的な行動を促進するとともに、ネガティブ行動を減少するこ

とを目的とした動物介在活動に着想を得たものである。ここでは、療育者による実時間での直観的なロボットの操作を可能にする制御手法を開発し、インタラクションに関する人の知見を可能な限り反映させた随伴性と一貫性のあるロボットとASD児の社会的相互作用を実現した。また、インタラクション中のASD児の快情動を推定するために装着型の表情計測デバイスを開発し、児童の自閉性傾向や現有スキルのバラつきを考慮した個人適応的なRAAを実現するための総合的なフレームワークの提案を行った。さらに、提案したロボットの制御手法によるロボット介在活動の検証実験を動物介在活動のポストセッションとして行い初回の反応を比較した結果、ロボット介在活動に参加したASD児のネガティブ行動の有意な減少が確認できたことから、提案手法である社会的な相互作用に基づくASD児の社会的行動の学習支援の可能性が確認できたといえる。

第5章 結論

本論文は、人間機械系の相互作用特性の理解に基づく行動学習支援に関して行った研究内容と、それによって得られた成果と知見をまとめたものである。ここでは、人や機械系の行動学習メカニズムに関する知見に基づき、行動学習支援に重要かつ環境やシステムに依存しない一般的な相互作用特性として随伴性と一貫性に着目したシステムの設計やインタラクションデザインを行った。

2章では、人と機械系の物理的な相互作用に基づく行動学習支援の事例として、駐車運転技量育成システムの開発に取り組んだ。ここでは、実車実験によって計測した熟練者の操作特性の理解に基づき、ステアリングを介した実時間での力覚呈示によって後ろ向き駐車に関するステアリング操作を支援するシステムを開発した。これにより、提案システムによる支援を行った被験者の後ろ向き駐車に関する運転技量が、支援を行わなかった被験者に比べて有意に向上したことを、開発したシミュレーション環境を用いた実験によって明らかにした。また、介入の強度やタイミング等の支援パラメータを一定とした場合支援に対する反応は被験者ごとに異なったことから、個々の運転者の既存の技量やスタイルに応じて、適応的に支援パラメータを調節する必要性も確認した。

3章では、人と機械系の情報的な相互作用に基づく行動学習支援の事例として、人が教示者としてロボットの行動学習を支援する手法について件キュを行った。これは、行動学習を行うロボットに対し人がオンラインで主観的な評価を与えることで行動学習に必要な評価関数を動的に作成・変更することを可能とする手法である。このとき、実時間での人の教示の随伴性と一貫性を仮定した起因の特定と誤り検出のアルゴリズムにより、人の教示意図を適切に推定して評価関数に反映させる手法も提案した。これにより、従来の機械系の行動学習手法における評価関

数の作成や最適化に関する設計者の作業負担を低減するとともに、既存の行動学習手法を評価関数を事前に最適化することが困難な実問題へ適用するための、人間機械系の相互作用からのアプローチを提案した。また、本手法を適用することで一般的な強化学習手法での行動学習効率が大幅に改善することを、被験者実験によって明らかにした。

4章では、人と機械系の社会的な相互作用に基づく行動学習支援の事例として、自閉症スペクトラム障がいを持つ児童の社会性向上やソーシャルスキルの学習支援を目的とする、ロボット介在活動について取り組んだ。ここでは、動物との触れ合いを通じた快情動の誘発に基づく、自閉症児の自発的な社会的行動の促進を目指す動物介在活動に着想を得て、ロボットとの自由なインタラクションを通じた自閉症児主体での療育手法を提案した。このとき、複雑多自由度な人型ロボットの行動を療育者が実時間で任意に操作するための人形型インタフェースを用いた制御手法を開発し、個々の児童の特性やロボットに対する反応に応じた即興的な社会的インタラクションを実現した。これにより、動物介在活動と比較したロボット介在活動の特徴や、それによる自閉症児の社会的行動の学習を支援出来る可能性を示唆する結果が実験によって得られた。今後は実験結果についてのより詳細な分析を行い、自閉症児の表情とロボットの行動やインタラクションの文脈との関係を調べることで、快情動を誘発するロボットの行動特性を児童ごとに比較し個人適応的な介入の必要性を明らかにする。また、得られた結果に基づき各児童に対して適切な応答を自律的に行うロボットの行動プログラムの開発や、装着型デバイスによって計測した表情を実時間フィードバックとしてロボットの行動学習に用いることも検討している [19]。

以上の事例研究を通して得られた知見に基づき、今後は人と機械システムの長期継続的な相互作用を通じて、人の主観性や個人差をシステムの制御に積極的に取り入れるための研究を行っていきたいと考えている。人は人工物とは異なり意思を持ち多様性があるとともに、身体的・生理的・社会的特性のバラつきと変化を持つ。しかしながら、このような性質はシステムの設計時において扱いが困難であるため、一般に系内に人を含む機械システムは人の主観や個人差を排した普

遍的かつ時不変な特性のみを考慮した設計・仕様となることが多い。本研究の事例研究においても、例えば第3章で人の教示入力にかかる時間遅れを計測し、その上限と下限の値を教示意図を推定するための時定数として用いたが、年齢性別の異なる被験者で同様の数値が得られるとは限らない。また、ロボットのどのような行動を重視して報酬を与えるかという、被験者ごとの主観的価値に基づくロボットへの教示方略の違いも実験結果から見て取れる。さらに、第2章で述べた運転技量育成システムにおいても、支援パラメータを運転者の既存の技量や腕力などの個人差を考慮せず一定としていたため、被験者によって行動学習の支援効果に違いが生じるという結果が得られた。これらは、より効率的かつ人と機械システムの相互作用に基づく行動学習支援の実現のために、人の主観性や個人差をシステムが理解し動作に反映させることの重要性を示唆している。このことから、第3章においては個々の児童の特性に基づく個人適応的なロボットの遠隔操作を実現するための手法を開発したが、社会的インタラクションの文脈に沿った適切な行動の選定は療育者自身の経験に依存している。また、ロボットの操作に一人の療育者が専念する必要があるなど、人の作業負担の観点からも継続的な相互作用を通じて個々の児童ごとに適切な応答をロボットが自律的に学習するための仕組みが重要である。これを実現するためのアプローチとして、第3章で提案した人からの主観的なフィードバックに基づくロボットの行動学習支援手法を用いることが考えられる。ここでは、装着型インタフェースを用いて計測した表情を実時間フィードバックとしてロボットの学習に活用することを考えている。それにより、個々の児童の快情動の生起に関連する行動をロボットが自律的に学習することができるとともに、社会的な文脈における表情が持つ意味や意義について自閉症児の自覚を促すことで、意識的に社会的ポジティブ行動を実行する動機付けにもなり得ると考えられる。本研究では、このように学習能力を持つ人と機械システムが相互作用によって互いの学習過程に介入し、それを促進するような正の学習ループ系を人間機械系の相互学習と定義する。今後は、この人間機械系の相互学習によって得られる恩恵として機械系の個人適応制御に着目し、生体電位信号や生理指標から推定した情動や社会的振る舞いなど、認知的・身体的負担の少

ない継続的かつ連続的な継続的に取得出来る情報をフィードバックとした機械系の行動学習メカニズムによって、人の好みや特性が直に反映された情報に基づく機械系の行動制御が期待出来る。これにより、人の主観的な感覚や印象と機械システムの動作を滑らかにつなぐ、次世代の人間機械協調システムの実現を目的として研究を行っていきたい。

参考文献

- [1] C. E. Billings. *Aviation automation: the search for a human-centered approach*. Human factors in transportation. Lawrence Erlbaum Associates Publishers, 1997.
- [2] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*, volume 28. MIT press, 1998.
- [3] A. Cypher and D. C. Halbert. *Watch what I Do: Programming by Demonstration*. MIT Press, 1993.
- [4] S. Schaal. Is imitation learning the route to humanoid robots? *Trends in cognitive sciences*, 3(6):233–242, 1999.
- [5] V. Kruger, D. Herzog, S. Baby, A. Ude, and D. Kragic. Learning actions from observations. *Robotics Automation Magazine, IEEE*, 17(2):30–43, 2010.
- [6] N. Delson and H. West. Robot programming by human demonstration: the use of human variation in identifying obstacle free trajectories. In *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*, pages 564–571 vol.1, 1994.
- [7] C. G. Atkeson and S. Schaal. Robot learning from demonstration. In Jr. Douglas H. Fisher, editor, *Machine Learning: Proceedings of the Fourteenth International Conference (ICML '97)*, pages 12–20. Morgan Kaufmann, San Francisco, CA, 1997.

-
- [8] C. G. Atkeson and S. Schaal. Learning tasks from a single demonstration. In *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on*, volume 2, pages 1706–1712. IEEE, 1997.
- [9] T. Inamura, I. Toshima, H. Tanie, and Y. Nakamura. Embodied symbol emergence based on mimesis theory. *The International Journal of Robotics Research*, 23(4-5):363–377, 2004.
- [10] 稲邑 哲也 and 中村 仁彦. ミメシス理論に基づく見まね学習とシンボル発達の統合モデル. 日本神経回路学会誌 = *The Brain and neural networks*, 12(1):74–80, mar 2005.
- [11] P. Kormushev, D. N. Nenchev, S. Calinon, and D. G. Caldwell. Upper-body kinesthetic teaching of a free-standing humanoid robot. In *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, pages 3970–3975, Shanghai, China, 2011.
- [12] D. H. Grollman and A. Billard. Donut as i do: Learning from failed demonstrations. In *ICRA*, pages 3804–3809. IEEE, 2011.
- [13] A. Lockerd and C. Breazeal. Tutelage and socially guided robot learning. In *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, volume 4, pages 3475–3480 vol.4, 2004.
- [14] M. Nakatani, K. Suzuki, and S. Hashimoto. Subjective-evaluation oriented teaching scheme for a biped humanoid robot. In *2003 IEEE-RAS International Conference on Humanoid Robots (Humanoids2003)*, 2003.
- [15] M. Riley, A. Ude, C. Atkeson, and G. Cheng. Coaching: An approach to efficiently and intuitively create humanoid robot behaviors. In *Humanoid*

-
- Robots, 2006 6th IEEE-RAS International Conference on*, pages 567–574, 2006.
- [16] M. Hirokawa and K. Suzuki. Coaching to enhance the online behavior learning of a robotic agent. In Rossitza Setchi, Ivan Jordanov, Robert J. Howlett, and Lakhmi C. Jain, editors, *Knowledge-Based and Intelligent Information and Engineering Systems*, volume 6276 of *Lecture Notes in Computer Science*, pages 148–157. Springer Berlin Heidelberg, 2010.
- [17] M. Hirkoawa and K. Suzuki. Coaching robots: Online behavior learning from human subjective feedback. In Ivan Jordanov and Lakhmi C Jain, editors, *Innovations in Intelligent Machines -3*, volume 442 of *Studies in Computational Intelligence*, pages 37–51. Springer Berlin Heidelberg, 2013.
- [18] H. Takagi. Interactive evolutionary computation: fusion of the capabilities of ec optimization and human evaluation. *Proceedings of the IEEE*, 89(9):1275–1296, 2001.
- [19] A. Gruebler, V. Berenz, and K. Suzuki. Coaching robot behavior using continuous physiological affective feedback. In *Humanoid Robots (Humanoids), 2011 11th IEEE-RAS International Conference on*, pages 466–471, 2011.
- [20] M. Aisen, H. Krieb, N. Hogan, F. McDowell, and B. T. Volpe. The effect of robot-assisted therapy and rehabilitative training on motor recovery following stroke. *Archives of Neurology*, 54(4):443–446, 1997.
- [21] G. S. Sawicki, K. E. Gordon, and D. P. Ferris. Powered lower limb orthoses: applications in motor adaptation and rehabilitation. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pages 206–211, 2005.

-
- [22] Y. Sankai. Leading edge of cybernics: Robot suit HAL. In *SICE-ICASE, 2006. International Joint Conference*, pages P-1 – P-2, oct. 2006.
- [23] Y. Sankai. HAL: Hybrid assistive limb based on cybernics. In Makoto Kaneko and Yoshihiko Nakamura, editors, *Robotics Research*, volume 66 of *Springer Tracts in Advanced Robotics*, pages 25–34. Springer Berlin / Heidelberg, 2011.
- [24] K. Yamawaki, R. Ariyasu, S. Kubota, H. Kawamoto, Y. Nakata, K. Kamibayashi, Y. Sankai, K. Eguchi, and N. Ochiai. Application of robot suit hal to gait rehabilitation of stroke patients: A case study. In Klaus Miesenberger, Arthur Karshmer, Petr Penaz, and Wolfgang Zagler, editors, *Computers Helping People with Special Needs*, volume 7383 of *Lecture Notes in Computer Science*, pages 184–187. Springer Berlin Heidelberg, 2012.
- [25] 高田 洋平, 佐伯 碧, 井澤 淳, 武田 湖太郎, 大高 洋平, 伊藤宏司, and 近藤 敏之. 視覚・体性感覚刺激と運動意図が事象関連脱同期に与える影響. 電子情報通信学会技術研究報告. *MBE, MEとバイオサイバネティクス*, 111(482):71–76, mar 2012.
- [26] D. J. Ricks and M. B. Colton. Trends and considerations in robot-assisted autism therapy. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 4354–4359, 2010.
- [27] S. Thill, C. A. Pop, T. Belpaeme, T. Ziemke, and B. Vanderborght. Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook. *Paladyn*, 3(4):209–217, 2012.
- [28] G. Pioggia, M. L. Sica, M. Ferro, R. Iglizzi, F. Muratori, A. Ahluwalia, and D. De Rossi. Human-robot interaction in autism: Face, an android-based social therapy. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pages 605–612, 2007.

-
- [29] B. F. Skinner. The behavior of organisms: An experimental analysis. 1938.
- [30] J. Bowlby. Attachment and loss:volume 1:attachment. *The International Psycho-Analytical Library, The Hogarth Press*, 1969.
- [31] J.-M. Hoc, M. S. Young, and J.-M. Blosseville. Cooperation between drivers and automation: implications for safety. *Theoretical Issues in Ergonomics Science*, 10(2):135–160, 2009.
- [32] J. Sil and T. Srikanthan. A hybrid approach to designing an autonomous driving alert system using geometrical features and gray level information of face images. In *IEEE Region 5, 2003 Annual Technical Conference*, pages 33–38, 2003.
- [33] N. Kaempchen, B. Schiele, and K. Dietmayer. Situation assessment of an autonomous emergency brake for arbitrary vehicle-to-vehicle collision scenarios. *Intelligent Transportation Systems, IEEE Transactions on*, 10(4):678–687, 2009.
- [34] T. Imkamon, P. Saensom, P. Tangamchit, and P. Pongpaibool. Detection of hazardous driving behavior using fuzzy logic. In *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008. ECTI-CON 2008. 5th International Conference on*, volume 2, pages 657–660, 2008.
- [35] T. Inagaki. Design of humanmachine interactions in light of domain-dependence of human-centered automation. *Cognition, Technology and Work*, 8(3):161–167, 2006.
- [36] M. R. Endsley and E. O. Kiris. The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(2):381–394, 1995.

-
- [37] P. G. Griffiths and R. B. Gillespie. Sharing control between humans and automation using haptic interface: primary and secondary task performance benefits. *Human Factors*, 47(3):574–590, 2005.
- [38] X.-D. Yang, W. F. Bischof, and P. Boulanger. Validating the performance of haptic motor skill training. In *Haptic interfaces for virtual environment and teleoperator systems, 2008. haptics 2008. symposium on*, pages 129–135, 2008.
- [39] S. Azzi, G. Reymond, F. Merienne, and A. Kemeny. Eco-driving performance assesment with in-car visual and haptic feedback assistance. *Journal of Computing and Infomation Science in Engineering*, 11(4), 2011.
- [40] E. H. F. van Asseldonk, M. Wessels, A. H. A. Stienen, F. C.T. van der Helm, and H. van der Kooij. Influence of haptic guidance in learning a novel visuo-motor task. *Journal of Physiology-Paris*, 103(35):276 – 285, 2009.
- [41] 沖田 齡次, 柴川 学, 佃 厚典, 石井 洋輔, 伊藤 剛豊, and 川原直人. インテリジェント ドライブ マスタ (i-dm) の概要. マツダ技報, 21, 2011.
- [42] 加藤 晋, 美濃部 直子, 川合 真弓, and 津川 定之. ビジュアルガイドによる運転支援システム : Hmi 及び協調走行の提案 (its・車, its 画像処理, 映像メディア及び一般). 電子情報通信学会技術研究報告. *ITS*, 104(647):13–18, jan 2005.
- [43] M. Tonnis, C. Lange, and G. Klinker. Visual longitudinal and lateral driving assistance in the head-up display of cars. In *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, pages 91–94, 2007.
- [44] 酒井 和彦. 世界初アラウンドビューモニター. 自動車技術 = *Journal of Society of Automotive Engineers of Japan*, 62(3):100–101, mar 2008.

-
- [45] N. Memarovic. *The Influence of Personal Navigation Devices on Drivers' Visual Attention on the Road Ahead and Driving Performance*. University of New Hampshire (Department of Electrical and Computer Engineering), 2009.
- [46] 牧野 靖, 田中 優, 遠藤 知彦, and 川端 佑輝子. インテリジェントパーキングアシストの開発. *自動車技術 = Journal of Society of Automotive Engineers of Japan*, 60(10):47–52, oct 2006.
- [47] 酒井 克博. スマートパーキングアシストシステムの開発. *自動車技術*, 63(12):60–64, dec 2009.
- [48] 安部 原也, 伊藤 誠, and 田中 健次. 自動化システムに対する信頼感の推移: 誤動作発生パターンへの依存性. *計測自動制御学会論文集*, 36(12):1138–1144, dec 2000.
- [49] J. Navarro, F. Mars, and M. S. Young. Lateral control assistance in car driving: classification, review and future prospects. *Intelligent Transport Systems, IET*, 5(3):207–220, 2011.
- [50] L. Saleh, P. Chevrel, F. Mars, J.-F. Lafay, and F. Claveau. Human-like cybernetic driver model for lane keeping. *World Congress*, 18(1):4368–4373, 2011.
- [51] L. Saleh, P. Chevrel, F. Claveau, J.-F. Lafay, and F. Mars. Shared steering control between a driver and an automation: Stability in the presence of driver behavior uncertainty. *Intelligent Transportation Systems, IEEE Transactions on*, 14(2):974–983, 2013.
- [52] M. Deroo, J.-M. Hoc, and F. Mars. Effect of strength and direction of haptic cueing on steering control during near lane departure. *Transportation Research Part F: Traffic Psychology and Behaviour*, 16(0):92 – 103, 2013.

-
- [53] M. Mulder, D. A. Abbink, and E. R. Boer. The effect of haptic guidance on curve negotiation behavior of young, experienced drivers. In *Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on*, pages 804–809, 2008.
- [54] N. Kuge and T. Yamamura. Study on driver’s car following abilities based on an active haptic support function. *SAE Transactions Journal of Passenger Cars - Electronic and Electrical Systems*, 115(7):157–169, 2006.
- [55] D. Powell and M.K. O’Malley. The task-dependent efficacy of shared-control haptic guidance paradigms. *Haptics, IEEE Transactions on*, 5(3):208–219, 2012.
- [56] 山門 誠, 高橋 絢也, and 齋藤 真二郎. 安全走行を支援する新しい車両運動制御技術「g-vectoring 制御」(特集環境, 安全, 快適を実現するオートモティブシステム開発技術). *日立評論*, 91(10):784–787, oct 2009.
- [57] T. Ohtsuka and H. A. Fujii. Real-time optimization algorithm for nonlinear receding-horizon control. *Automatica*, 33(6):1147 – 1154, 1997.
- [58] T. Kobayashi and S. Majima. Real-time optimization control for parking a vehicle automatically. In *Vehicle Electronics Conference, 2001. IVEC 2001. Proceedings of the IEEE International*, pages 97–102, 2001.
- [59] J. J. Grefenstette. Credit assignment in rule discovery systems based on genetic algorithms. *Machine Learning*, 3(2-3):225–245, 1988.
- [60] 前田 陽一郎 and 花香 敏. Shaping 強化学習を用いた自律エージェントの行動獲得支援手法. *知能と情報 : 日本知能情報フアジィ学会誌 : journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, 21(5):722–733, oct 2009.

-
- [61] K. Doya. Temporal difference learning in continuous time and space. In *Advances in Neural Information Processing Systems 8*, pages 1073–1079. MIT Press, 1996.
- [62] K. Doya. Reinforcement learning in continuous time and space. *Neural Computation*, 12:219–245, 2000.
- [63] L. Wing. Autistic spectrum disorders. *BMJ: British Medical Journal*, 312(7027):327, 1996.
- [64] Y. S. Kim, B. L. Leventhal, Y.-J. Koh, E. Fombonne, E. Laska, E.-C. Lim, K.-A. Cheon, S.-J. Kim, Y.-K. Kim, H.-K. Lee, et al. Prevalence of autism spectrum disorders in a total population sample. *American Journal of Psychiatry*, 168(9):904–912, 2011.
- [65] J. Baio. Prevalence of autism spectrum disorders: Autism and developmental disabilities monitoring network, 14 sites, united states, 2008. morbidity and mortality weekly report. surveillance summaries. volume 61, number 3. *Centers for Disease Control and Prevention*, 2012.
- [66] S. J Rogers and L. A. Vismara. Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child & Adolescent Psychology*, 37(1):8–38, 2008.
- [67] B. Bettelheim. *The empty fortress: infantile autism and the birth of the self*. Free Press, 1967.
- [68] E. T. Bekele, U. Lahiri, A. R. Swanson, J. A. Crittendon, Z.E. Warren, and N. Sarkar. A step towards developing adaptive robot-mediated intervention architecture (aria) for children with autism. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 21(2):289–299, 2013.

-
- [69] D. J. Ricks and M. B. Colton. Trends and considerations in robot-assisted autism therapy. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 4354–4359, 2010.
- [70] G. Sugai, R. H. Horner, G. Dunlap, M. Hieneman, T. J Lewis, C. M. Nelson, T. Scott, C. Liaupsin, W. Sailor, A. P. Turnbull, et al. Applying positive behavior support and functional behavioral assessment in schools. *Journal of Positive Behavior Interventions*, 2(3):131–143, 2000.
- [71] F. Martin and J. Farnum. Animal-assisted therapy for children with pervasive developmental disorders. *Western Journal of Nursing Research*, 24(6):657–670, 2002.
- [72] B. M. Prizant. *The SCERTS Manual: A Comprehensive Educational Approach for Young Children with Autism Spectrum Disorders*. Paul H. Brookes, 2005.
- [73] J. Nimer and B. Lundahl. Animal-assisted therapy: A meta-analysis. *Anthrozoos: A Multidisciplinary Journal of The Interactions of People and Animals*, 20(3):225–238, 2007.
- [74] H. Keino, A. Funahashi, H. Keino, C. Miwa, M. Hosokawa, Y. Hayashi, and K. Kawakita. Psycho-educational horseback riding to facilitate communication ability of children with pervasive developmental disorders. *Journal of equine science*, 20(4):79–88, dec 2009.
- [75] A. Funahashi, A. Gruebler, T. Aoki, H. Kadone, and K. Suzuki. Brief report: The smiles of a child with autism spectrum disorder during an animal-assisted activity may facilitate social positive behaviors quantitative analysis with smile-detecting interface. *Journal of Autism and Developmental Disorders*, pages 1–9, 2013.

-
- [76] D. Feil-Seifer and M. Mataric. Robot-assisted therapy for children with autism spectrum disorders. In *Proceedings of the 7th international conference on Interaction design and children*, pages 49–52. ACM, 2008.
- [77] E. Barakova and T. Lourens. Interplay between natural and artificial intelligence in training autistic children with robots. In JosManuel Ferrndez Vicente, JosRamn lvarez Snchez, Flix Paz Lpez, and Fco.Javier Toledo Moreo, editors, *Natural and Artificial Models in Computation and Biology*, volume 7930 of *Lecture Notes in Computer Science*, pages 161–170. Springer Berlin Heidelberg, 2013.
- [78] H. Kozima and C. Nakagawa. Interactive robots as facilitators of children ’ s social development. *Mobile robots: Toward new applications*, pages 269–286, 2006.
- [79] L. Boccanfuso and J. M. O ’ Kane. Adaptive robot design with hand and face tracking for use in autism therapy. In *Social Robotics*, pages 265–274. Springer, 2010.
- [80] D. Mazzei, A. Greco, N. Lazzeri, A. Zaraki, A. Lanata, R. Iglizzo, A. Mancini, F. Stoppa, E.P. Scilingo, F. Muratori, and D. De Rossi. Robotic social therapy on children with autism: Preliminary evaluation through multi-parametric analysis. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*, pages 955–960, 2012.
- [81] T. Sugihara and Y. Nakamura. Whole-body cooperative balancing of humanoid robot using cog jacobian. In *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, volume 3, pages 2575–2580 vol.3, 2002.

- [82] T. Sugihara and Y. Nakamura. Whole-body cooperative cog control through zmp manipulation for humanoid robots. In *2nd Int. Symp. on Adaptive Motion of Animals and Machines (AMAM2003)*, 2003.
- [83] A. Gruebler and K. Suzuki. Measurement of distal emg signals using a wearable device for reading facial expressions. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 4594–4597, 2010.
- [84] F. R. Volkmar and L. C. Mayes. Gaze behavior in autism. *Development and Psychopathology*, 2:61–69, 1 1990.

謝辞

本研究の遂行ならびに論文の作成にあたり，懇切なる御指導を賜りました筑波大学大学院システム情報工学研究科准教授 鈴木健嗣先生 に謹んで感謝の意を表します。

本研究の遂行にあたり，共同研究者として有益な御助言と御教示を賜りました愛知県心身障害者コロニー発達障害研究所教育福祉学部発達教育研究室室長 舟橋厚先生 に心より謝意を申し上げます。

本論文をまとめるにあたり，有益な御助言と御教示を賜りました筑波大学大学院システム情報工学研究科教授 山海嘉之先生，同研究科准教授 長谷川泰久先生，伊藤誠先生 に心より謝意を申し上げます。

本研究の遂行ならびに論文の作成にあたり御協力いただいた，筑波大学大学院システム情報工学研究科人工知能研究室の皆様へ感謝致します。また，諸般の事務手続きを行っていただいた筑波大学大学院システム情報工学研究科人工知能研究室秘書 安田敦子氏，栗本千加氏，入江夕子氏，大木美加氏，研究補佐員 倉本尚美氏，佐藤久美子氏 に感謝いたします。

最後に，長い間遠方から支えてくれた家族に，心から感謝の意を表します。

2014年3月 廣川 暢一

研究業績

主要論文

- [1] 廣川 暢一, 鈴木 健嗣, 「教示者による学習支援に基づくエージェントのオンライン行動学習」, 人工知能学会論文誌, vol. 25, no. 6, pp.694-702, 2010.
- [2] Hirokawa, M., Suzuki, K., “ Coaching Robots: Online Behavior Learning from Human Subjective Feedback ”, Studies in Computational Intelligence, vol. 442, pp. 37-51, 2013.
- [3] 廣川暢一, 上杉直久, 古郡了, 北川朋子, 鈴木健嗣, 「運転技量育成のための力覚呈示用いた運転行動支援システム」, 計測自動制御学会論文誌, vol. 49, no. 6, 2013.

国際会議

- [1] Hirokawa, M., Suzuki, K., “ Coaching to Enhance the Online Behavior Learning of a Robotic Agent ”, 14th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, Cardiff, UK, 2010.
- [2] Hirokawa, M., Uesugi, N., Furugori, S., Kitagawa, T., Suzuki, K., “ A Haptic Instruction based Assisted Driving System for Training the Reverse Parking ”, Proceedings of IEEE International Conference on Robotics and Automations, pp. 3713-1718, 2012.
- [3] Hirokawa, M., Funahashi, A., Suzuki, K., “ On the Analysis of Eye-Contact of

- Children with Autism Spectrum Disorders in Robot-Assisted Activities”, 8th ACM/IEEE International Conference on Human-Robot Interaction, 2013.
- [4] Hirokawa, M., Funahashi, A., Suzuki, K., “A Doll-type Interface for Real-Time Humanoid Teleoperation in Robot-Assisted Activity”, 9th ACM/IEEE International Conference on Human-Robot Interaction, 2014.

国内会議

- [1] 廣川暢一, 鈴木健嗣, 「コーチングによる報酬関数の動的生成に基づくエージェントのオンライン行動学習」, HAI シンポジウム 2009, 東京, 2009.
- [2] 廣川暢一, 鈴木健嗣, 「コーチングに基づくロボットのオンライン行動学習」, 日本機械学会ロボティクス・メカトロニクス講演会, 福岡, 2009.
- [3] 廣川暢一, 鈴木健嗣, 「透過型デバイスを用いたコーチングによるロボットの学習支援」, 日本機械学会ロボティクス・メカトロニクス講演会, 旭川, 2010.
- [4] 上杉直久, 廣川暢一, 北川朋子, 古郡了, 鈴木健嗣, 「駐車運転の技量向上を促進する運転支援システムの研究」, 自動車技術会関東支部学術研究講演会, 東京, 2012.
- [5] 廣川暢一, 鈴木健嗣, 「身体表象と実像を関連づける人形型インタフェースによる遠隔ロボット操作」, 日本機械学会ロボティクス・メカトロニクス講演会, 浜松, 2012.
- [6] 廣川暢一, 鈴木健嗣, 「ステアリングへの力覚呈示を介した駐車技量育成システム」, 日本ロボット学会第30回学術講演会, 札幌, 2012.
- [7] 廣川暢一, 舟橋厚, 鈴木健嗣, 「発達障がい児の社会的行動を促進する療育支援ロボットのための行動制御」, 日本機械学会ロボティクス・メカトロニクス講演会, つくば, 2013.
- [8] Hirokawa, M., Funahashi, A., Suzuki, K., “On the Design of Robot-Assisted Activities based on Animal-Assisted Activities for Children with Autism Spectrum Disorders”, The 36th Annual Meeting of the Japan Neuroscience Society,

Kyoto, 2013.

- [9] Funahashi, A., Gruebler, A., Aoki, T., Hirokawa, M., Kadone, H., Suzuki, K., “Relationship between social behaviors and smiles of a child with Autism Spectrum Disorders during the animal-assisted or robot-assisted activities - behavioral observation and quantitative analysis using a smile detection interface- ”, The 36th Annual Meeting of the Japan Neuroscience Society, Kyoto, 2013.

受賞等

- [1] 2012年度自動車技術会関東支部学術研究講演会ベストペーパー賞, 2012.
[2] 計測自動制御学会システム・情報部門技術賞, 2013.