

生活支援ロボットによる
時空間遮蔽近傍情報を用いた作用推定
に基づく道具利用タスク学習

長濱 虎太郎

目次

第1章	序論	1
1.1	本研究の背景と目的	3
1.2	本論文の構成	7
第2章	作用・機能推定に基づく道具利用タスク学習	9
2.1	はじめに	11
2.2	道具を用いたタスクの観察学習	11
2.3	ロボットの道具を用いた作業における問題点	18
2.4	視覚重畳関係を利用した作用推定	25
2.4.1	作用推定と対象の遮蔽	25
2.4.2	視覚重畳関係とその種類	28
2.5	重畳関係を推定するための手がかり：空間遮蔽近傍情報と遮蔽時刻近傍情報	38
2.6	検証に用いるタスク学習システム	40
2.7	道具利用タスク学習における用語	46
2.8	おわりに	47
第3章	道具を用いた生活支援タスクを学習するロボットシステム	49
3.1	はじめに	51
3.2	日用道具を用いて生活支援をおこなうロボット	51
3.3	日用道具を用いて生活支援をおこなうロボットの構成法	52
3.3.1	HRP-2VZのハードウェア構成	53
3.3.2	ソフトウェア構成と道具利用のための動作生成法	53
3.4	人間と同等の視力を有するロボットヘッドの構成法	56
3.4.1	ズームヘッドの開発史	56
3.4.2	ズームヘッドのハードウェア構成	57
3.4.3	ズームヘッドのソフトウェア構成	58
3.4.4	広視野視覚と高視力視覚を併用した注視制御	63
3.4.5	輻輳ズームステレオ機構に関する考察	67
3.4.6	ズームヘッド改良版“CV-Head Mk.II”	68
3.4.7	CV-Head Mk.IIのハードウェア・ソフトウェア構成	69

3.5	HRP-2VZによる道具利用動作の考察	71
3.6	おわりに	82
第4章	空間遮蔽近傍情報を用いた見えない道具の状態推定	83
4.1	はじめに	85
4.2	一部が見えない道具や操作対象の状態推定のための視覚機能	88
4.3	遮蔽関係のある複数平面領域の追跡法	88
4.3.1	粒子フィルタを用いた平面道具追跡	88
4.3.2	尤度計算法	88
4.3.3	非遮蔽予測領域を用いた対象領域の追跡法	90
4.3.4	非遮蔽予測領域を用いた複数対象の追跡・基礎実験	90
4.4	重畳関係と随伴性の評価法と作用推定	91
4.4.1	作用推定処理の流れ	93
4.4.2	重畳関係の分類法	95
4.4.3	運動随伴性の評価法	96
4.5	重畳関係・随伴性を用いた作用分類実験	97
4.5.1	作用分類のシミュレーション実験	97
4.5.2	実画像を用いた実験の条件	98
4.5.3	実画像を用いた実験の結果と評価	102
4.6	ロボットによる見えない対象の状態推定に基づく道具利用	108
4.6.1	重畳関係を用いた状態推定と動作修正	108
4.6.2	自らが操る道具を追跡するための対象動作モデル	112
4.7	見えない対象の状態推定に基づく道具利用実験	112
4.7.1	実験1：ファイルのフォルダ挿入実験	112
4.7.2	実験2：目玉焼きのよそいつけ作業	119
4.8	おわりに	125
第5章	遮蔽時刻近傍情報を用いた見えない作用推定に基づく学習法	127
5.1	はじめに	129
5.2	視覚重畳関係を利用するV-FESの構成	131
5.3	遮蔽時刻前後の特徴を用いた作用推定	135

5.3.1	作用推定のためのメタ情報	135
5.3.2	遮蔽が発生するシミュレーション実験の入力データ	136
5.3.3	追跡結果からの運動情報推定	138
5.3.4	知識 KNO-1 から KNO-3 の利用	144
5.3.5	運動随伴性の評価法	147
5.3.6	知識 KNO-4 と KNO-5 の利用	148
5.4	時系列メタ情報からの作用・機能推定	151
5.5	ロボットによる実観察画像列を用いた観察学習実験	154
5.5.1	柔軟物を含むタスク観察学習法	154
5.5.2	実観察画像列を用いたタスク学習結果	156
5.6	LINE-MOD を用いたテクスチャレス物体追跡と観察学習	168
5.6.1	LINE-MOD	168
5.6.2	LINE-MOD の物体認識特性に関する考察	168
5.6.3	ロボットによる実環境での片付け観察実験	172
5.6.4	V-FES を用いた作用・道具機能推定結果	177
5.7	おわりに	183
第 6 章	作用・機能推定結果に基づくロボットの身体に合わせたタスク再現	187
6.1	はじめに	189
6.2	道具を利用したタスクの自律計画法	189
6.2.1	PDDL によるタスク記述と道具利用法表現	190
6.3	道具を用いたタスクの再現シミュレーション実験	199
6.3.1	単腕ロボット HSR でのタスク再現実験	199
6.3.2	双腕ロボット HRP-2VZ でのタスク再現実験	202
6.4	視覚重畳関係を用いる観察学習法の日常タスクにおける利用	211
6.4.1	収納に関する観察学習とその再現	211
6.4.2	洗濯物の分類に関する観察学習とその再現	219
6.5	おわりに	233
第 7 章	結論	235
7.1	各章のまとめと結論	237

7.2 将来の展望	240
謝辞	243
発表文献等	248
参考文献	258

第1章

序論

1.1 本研究の背景と目的

本研究で目標とするのは、家の中で人の傍にいて、人の手助けをしてくれるロボットである。

近年、人間に類似した体の構造を有するロボットの開発が、大学や研究機関、企業等でおこなわれている。本田技研工業による ASIMO[33] やトヨタ自動車によるパートナーロボット [21]、そして産業技術総合研究所による HRP-4C[98] 等が、博物館の展示物説明やステージでの楽器演奏、MC 等のエンターテイメント分野で活用されていることは、一般にも広く知られている事実である。一方で、人間に類似した形のロボットは、人間が普段利用している環境をそのまま用いて、人間の作業を肩代わりしていく存在となり得るとも考えられる。このような期待から、研究用人型ロボットハードウェアとして HRP-2[41]、早稲田大学による TWENDY-ONE[38]、米国 Willow Garage 社による PR2[26, 74]、イタリア IIT による iCub[48]、ドイツ KIT による ARMAR[3] 等が開発されるとともに、OpenRTM[2]、ROS[62] のようなロボット用ミドルウェアや OpenCV[12] のような画像処理ソフトウェアライブラリが整備され、実環境でのロボットの認識・行動システムの研究開発が進められている。

このような中、ロボットを少子高齢化社会を支える技術として、家庭内で活用していくことも期待されている [95]。2013 年現在で既に家庭内へ導入されているロボットとして、iRobot 社の Roomba[36] に代表される掃除ロボットが挙げられるが、より複雑な家事介護支援の課題として、食事の準備や片付け、洗濯をはじめとする衣類の操作等が残されている。これらの仕事をロボットが人の代わりにおこなうことができれば、高齢者や介護者等の負担を軽減するとともに、QOL (Quality of life) の向上へ貢献できると考えられる [85]。また、要介護者が家事等を無理なくロボットと一緒にこなうことができれば、レクリエーションとして基礎体力の向上に役立つ [125] 他、精神的な支援へもつながると考えられる。

家庭内で家事支援をおこなうロボットは世界各国で、各操作のための認識手法や動作生成法に関する研究開発が盛んにおこなわれている [3, 9, 23, 56, 71, 80] が、ロボットが各家庭へ入っていく際には、各家庭に固有の家事のやり方をロボット自身が学んでいく能力も必要となると考えられる。例えば、片付けの手伝いでは「何をどこに片付けるか」、料理の手伝いでは「どういう手順でどう作るか」、掃除の手伝

いでは「どの道具をどこに使うか」などは各家庭に固有の問題となり、工場ではインプットできない情報である。

このような情報を家庭内でエンドユーザがロボットへ伝える手法を考えると、以下のようなものが存在する。

プログラミングによる指示

現在、ほぼ全ての家事支援ロボットの実証研究でおこなわれている手法である ([80] など)。利点としてはタスクの指示ミスが少ないことが考えられ、ロボットの身体能力および指示者のプログラミング能力が十分であるならば、理想的にはロボットが可能な全てのタスクを指示可能であると考えられる。しかし、指示者がロボットのソフトウェアシステムに精通している必要があり、エンドユーザにこのような知識と能力を求めるのは難しい。

簡易プログラミングによる指示

GUI (graphical user interface) 等を用いることで、ロボットの動作・行動を簡易的に指示・設計していく手法も研究されており [66] , 趣味・教育用途などでは普及している例もある (教育版レゴ・マインドストーム NXT [89] 等) 。この手法の利点としては、プログラミングによる指示と同様に指示ミスが少ないこと、そして比較的 IT 技術に精通したエンドユーザならば、簡単なタスクであれば簡潔な指示が可能なことである。本研究で対象とするような高齢者等の支援においては、やはり操作方法の伝達の煩雑さが問題となると考えられるが、他の手法と併用することは、指示タスクの変更等に有用であると考えられる。ただし、人がロボットと一緒に作業をしながら指示することは不可能である。

コントローラを用いた直接教示・簡易指示

タブレット端末等を用いて簡易的にロボットへ動作指示を出し、実際にロボットを動かしながらタスク教示をおこなっていく手法である ([7, 145] 等)。この手法では、実環境中でロボットを直接動かすため、各状況に応じた教示をおこないやすいと考えられる、また、ロボットの視覚画像を利用することで、一人称視点での操作が可能であることなども利点として挙げられる。物を拾って持ってくる等の比較的簡易かつ定まった動作の指示であれば、動作に必要なパラメタを制限することで、実証試験レベルへ達している例 [115] も存在する。

しかし簡易プログラミングによる指示と同様、複雑な作業の指示を可能とするためには、複雑な操作方法の習熟が必要となると考えられる。また、人がロボットと一緒に作業をしながら指示することが難しいという欠点もある。さらに、ロボットが再利用可能な形で、タスクに関する知識を取得・蓄積するためには、単なるマニピュレータの軌跡レベル等に留まらない、そのタスクに真に必要な情報を自律的に獲得する必要がある。これは後述の直接教示、あるいは視覚教示にも共通する課題と言える。

直接教示

古くは工場のロボットアームへの動作教示方法としておこなわれてきた手法であり、近年でも、教示時の操作対象の視覚認識と組み合わせることで、再利用可能な形でのタスク学習の研究が進められている [17]。教示 1 回目から成功する可能性が高いこと、そして人間にとっても、特に子供等への教示方法と類似していることから、エンドユーザにも抵抗がない手法であると考えられる。一方、大きなロボットへの動作教示には大きな動きと力が必要であり、また道具を持った状態のロボットへ接触することの危険性も指摘されている [84]。

視覚教示（観察学習・模倣学習）

ロボットアームへの動作教示方法として、直接教示の欠点を克服するため、あるいは人間の観察・模倣学習能力を解き明かすためにも研究が進められている方法である [44, 108]。教示者はロボットの前で普段通りに作業をおこなうだけで良く、基本的な身体機能を有するエンドユーザにとっては特別な訓練を必要としない教示方法である。また人間にとって、他者への教示方法と類似していることから、エンドユーザへ導入しやすい手法であると考えられる。

本研究で注目するのは、最後に述べた視覚教示（観察学習・模倣学習）手法である。図 1.1 に、生活支援ロボットへの家事タスク視覚教示のイメージ図を示す。この手法が確立できれば、前述の利点のみならず、人がおこなっているタスクをその場で手伝うことなども同様に可能となると考えられる。ただし一般的な人の行動の視覚教示には、人間の意図している教示部分を、どのようにロボットへ推定させるか、という問題があり、文献 [112] で詳しく整理されている。この問題の部分的な解決法としては、タスク毎に真に正しい環境の状態変化の認識方法と、その変化を起

こすために必要な動作プリミティブを決定して用いる手法が、従来より提案されている [68]。生活支援ロボットの場面にも、ピックアップブレース、中に入れる、回す、かき混ぜる等の動作プリミティブは多くのタスクで頻出することから、別途動作計画をおこなうための知識を有し、その組み合わせ方をユーザが指示できるようにすべきだと、本論文では考える。理由の一つは、このようにタスクを動作プリミティブ単位で物体操作を記述し記憶しておくことで、動作プリミティブそのものの学習あるいは動作計画法の改良を、別の問題として考えることができ、タスク学習結果の再利用性が向上することが利点として考えられることである。複雑な環境下でもロボットモデルや目標座標を与えれば、障害物回避を含めたマニピュレータの軌道レベルの計画を自律的におこなうシステムの開発等が実際におこなわれており [22]、生活支援ロボットがこのような機能を有しておくことは、ユーザが簡単な動作を一つ一つ教える煩雑さを低減するためにも必要であると考えられる。さらに、家庭内に入るロボットには、専門知識を有しない不特定多数のユーザが運用する際の安全性も大きな課題となっている [135]。動作速度の計画法や障害物への対処法などは、前述の軌道レベルでの課題となり、ロボットの開発者が責任を持つためには、プリミティブに関する“カプセル化”¹が必要だと考えられる。

このような枠組みにおいても依然残されているのは、人を観察する際に、各動作の出現と、その目的や手順を推定するための手法に関する課題である。従来より、物体のピックアップブレース [44] や液体の注ぎ動作 [108] 等における、人間が操作する対象の変化や、複数の対象間の関係の変化等を認識し再現するための手法が研究されている。しかし道具を用いた対象の操作では、従来のタスク視覚学習研究で取り上げられていなかったような、上下・包含・突入関係を変化させる操作が現れる。例えば、鍋へ蓋をする、袋へ物を収納する、包丁で物を切るといった操作である。さらに、一つの道具が複数の機能を担う場合が存在する。例えば、包丁を使った調理においては、野菜を切るだけでなく、包丁を直角に使って野菜を寄せ集める、包丁に物を載せて運ぶ等の操作が存在する。一方で、同じような機能を有する道具でも、対象物に関して制限が存在する場合がある。例えば、箒はゴミを動かし、へらでは料理を動かすことができるが、これらの組み合わせを逆にすることは許されない。さらに、この家庭のこの絵皿には物を載せて運んではいけないといったよう

¹オブジェクト指向において、オブジェクト内部のデータを隠蔽すること

な，道具の機能に関する制約も存在する．したがって道具を用いた家事支援の観察学習においては，タスクの手順，操作目的，道具の機能や道具と操作対象の組み合わせを，上下・包含・突入関係の変化を含むタスクの観察から推定する視覚的枠組みが重要であると考えられる．

一方で，見ようとしても観察対象が見えないような操作も存在する．例えば物を箱にしまうような状況では，しまわれた対象は箱に覆われて，どのような角度から見ようとしても見えない，つまり認識不可能な状況に陥る．しかし片付けなどの観察学習では，このような操作は頻出することから，遮蔽下の対象や道具に関しても，どのような状況にあるかを推定しながら，教示者の教えたいタスクを学んでいく能力が必要になると考えられる．このような見えない状態を推定しなければならない状況は，ロボットが道具を用いてタスクを再現する際にも発生する．

本論文では，このような道具を用いたタスクの視覚学習，あるいはロボットが道具を用いてタスクを再現する際に，道具と操作対象の間に遮蔽が発生している状況下でも，それらの遮蔽下での状態を推定することで，その間に生じている相互作用，あるいは道具の機能を推定するための視覚機能について述べる．本研究では，遮蔽下の状況を推定するために遮蔽されている部分の周辺の領域（空間遮蔽近傍情報），そして道具や対象が遮蔽にいたるまでの状況，あるいは遮蔽が解消される状況の特徴（遮蔽時刻近傍情報）を用いる．また，見えている部分が小さい，あるいは薄い道具と操作対象を観測するための，高視力ヒューマノイドヘッドの構成法についても述べる．

1.2 本論文の構成

本章序論では，本研究の背景，目的と本論文の構成について述べる．2章作用・機能推定に基づく道具利用タスク学習では，生活支援ロボットが道具を用いたタスクを観察学習する際の問題点について論じる．観察時のみならず，道具を用いたタスクを遂行する際にも遮蔽下にある道具と操作対象の状態を推定することの重要性について述べ，視覚重畳関係を用いた推定方法を提案する．3章道具を用いた生活支援タスクを学習するロボットシステムでは，道具と操作対象を高視力視覚で観測するためのロボットヘッド，ヘッドを搭載して実験に用いた等身大ヒューマノイド



図 1.1: 生活支援ロボットによる片付け観察学習のイメージ

ロボットの構成法について述べる．さらにロボットへ実際に道具を用いた作業をおこなわせ，道具利用タスクにおける問題点と解決法について考察する．4章空間遮蔽近傍情報を用いた見えない道具の状態推定では，遮蔽下にある道具や操作対象の状況を，遮蔽されている領域の周辺領域（空間遮蔽近傍情報）を用いて推定するための視覚処理手法について述べ，それを用いたロボットによるタスク観察結果，ロボットによる道具利用タスク遂行結果についても述べる．5章遮蔽時刻近傍情報を用いた見えない作用推定に基づく学習法では，箱に対象をしまうような，完全に見えなくなる観察対象の状況を，その前後の状況（遮蔽時刻近傍情報）を用いて推定するための手法“V-FES”について述べ，実環境によるロボットのタスク観察実験により評価する．6章作用・機能推定結果に基づくロボットの身体に合わせたタスク再現では，5章までで推定したタスク目的や道具の機能に従って，観察学習時と変化した環境から，ロボットの体に合わせてタスクを遂行するための手法と評価実験について述べ，7章結論でまとめる．

第2章

作用・機能推定に基づく道具利用タスク学習

2.1 はじめに

本論文では、道具を用いる日常タスクを学習しながらおこなっていく生活支援ロボットが、物体の遮蔽関係という視覚的な手がかりを利用して、対象の状態推定をおこない活用していくための方法論とその有用性について論じる。

本章でははじめに、道具を用いる家事支援タスクと道具を用いたタスクの観察学習における課題について述べ、近接しあう観察対象あるいは操作対象同士の状態推定をおこなうことの重要性について述べる。この問題について、遮蔽関係そのものを手がかりとして活用していくことが有用である可能性について考察し、視覚重畳関係を定義し、その推定方法についても議論する。

視覚重畳関係を基軸とした視覚的手がかりの、ロボットの家事タスク観察学習や、ロボット自身の家事タスク遂行における有用性を検証するために、本研究で解くべき課題について述べるとともに、本研究で検証に用いる観察学習システムを示す。

なお本章の記述は著者論文 [53, 54, 100] の内容を含む。

2.2 道具を用いたタスクの観察学習

人間は家庭内の調理や掃除等の家事において、道具を用いることで、複雑な作業を素手よりも効率的に、そして安全で衛生的に作業を実現している。生活支援ロボットへも同様に、人間が利用している道具を利用させることができれば、普段人間がおこなっている家事タスクを肩代わりさせていくことが可能となると考えられる。さらに前章で考察したように、このような道具を用いた家事支援タスクを、一般のユーザが家庭内で視覚教示できれば、人間と同じ環境や道具を用いたロボットによる作業を、各家庭に応じたものへ容易に変更可能とできるであろう。

改めて、家庭へ入る生活支援ロボットが、家庭内での作業を学習することを考える。もしこのロボットが、基本的なものの持ち方、運び方を知らず、そこからユーザが教示しなければならないようであれば、ユーザへの負担は大きい。技術面でも、物体把持を含む動作計画研究はロボット一般に利用できる手法として進められている(文献 [22] 等)。そこで、基本的な認識や動作のためのスキルは工場レベルでインプットされていることを期待すべきであると考えられる。前者としては、ロボットの存在する環境や物体の位置・姿勢等の状態、そして存在の可否を推定する技術と、そ

の手がかりを学ぶ技術，さらに物体同士の関係を推定・認識する技術が挙げられる．後者としては，操作物体同士の近接・接触を主とする関係の変化を起こすようなマニピュレータの動作計画技術が挙げられる．

しかしそうであっても，各家庭に固有となる家事タスクの目的や道具の使い方，つまり道具の動かし方や，どの物体をどのタスクへの道具として利用するかは，家庭に入るまでインプットするのは難しい．またタスクの内容を家庭内でも簡易に変更できる枠組みは必須である．前者の例を挙げると，片付けのときに何をどこへしまうか，料理のときに何を切ってから何を混ぜて，何を焼くか，というものである．後者の例としては，例えば布巾や皿を考える．洗い終わった食器を拭くための布巾がどれで，床を拭くための雑巾がどれかは，その場で教えたり，あるいは普段家族が使っているものを覚えることが必要かつ重要であろう．従来より，画像処理により [20]，あるいはロボット自身が操作しながら [136, 146]，道具の機能部位を自律的に発見する研究がおこなわれてきている．しかし，1) 箸をつかって料理を動かしてはいけない，2) 皿についても，高級な皿は傷付かないように一つ一つ運ぶ必要があるが，そうでない皿は重ねて効率的に運んで良い等，物理的に可能な道具としての利用法を制限するような，家庭内のルールを学んでいく能力は必要となる．従って，本稿で注目し学習すべき事項とするものは，以下の二点とする．

タスクの目的

操作対象をどのような状態へ変化させたいかという情報．どこの中にどの物品を片付けるか，どの対象をどのような切断面の関係となるように切断するかといった情報である．

道具の使い方

物体にどのような操作を加えて良いかという情報．どの道具をどのように用いて良いか（運ぶ，切る等），といった利用可能情報である．

次に，生活環境で利用されている道具について考察する．日常生活で道具が多く用いられるシーンとして，調理シーンがある．生活支援の一環として，ロボットに調理を手伝わせることも大きな目的の一つとなるが，その中で出現する対象同士の作用のさせ方は，やはり多様なものとなる．ここで調理道具の解説書 [144] に現れる，道具利用を表す動詞と動作に関する名詞を全て書きだしたものを図 2.1 に示す．



図 2.1: 調理道具の解説 [144] に現れる，作用を表す言葉の種類

多様な言葉が存在することが分かるが，注目すべきは上下・包含関係の変化が重要なもの，あるいは上下・包含関係が付随するもの，そして突入関係の変化や対象の変形を引き起こすものが多く，全体の約半数を占めていることである．この表は一回でも出現した単語を取り上げているものであるため，実際に調理作業中に出現する頻度を示すものではない．しかし，上下・包含関係の変化や突入関係の変化を観測することが，複数対象の間に現れる多様な作用を推定するために必要であると考えられる．

文献 [111] では，生活環境内での人間行動を，国際生活機能分類を用いて分類している．この中では，持ち上げる，物を置く，手の込んだ食事の調理の順に発生の多い活動とされている．前述の調理にかかわる対象の状態変化が重要であることだけでなく，持ち上げる，物を置くことも，その地点に対象が存在する場合には上下・

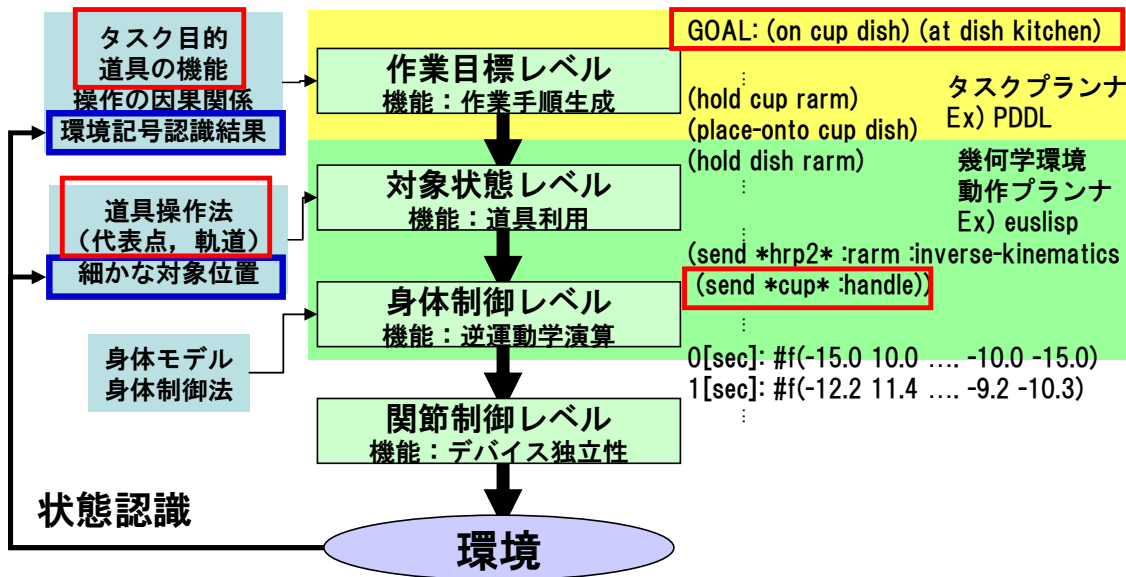


図 2.2: 生活支援ロボットの行動・動作記述レベル

包含関係の変化を伴うであろうことから、上下・包含関係の変化や突入関係の変化を観測することに重要性があると考えられる。

そして、ロボットへのタスクの視覚教示 [10] について考える。前章でも述べたとおり、ロボットへの物体操作を伴う作業の視覚教示の研究の歴史は工業用ロボットアームへのタスク視覚教示 [34, 44] にはじまり、エンドユーザのロボットへの作業指示手法として、コーディングが難しいロボット動作の自律獲得手法として、あるいは人間の模倣学習知能の解明の一貫として、物体操作だけでなく身体の動かし方 (ジェスチャの教示 [15], 舞踊動作の伝達 [121] 等) に関しても、研究がおこなわれている [10]。文献 [112] では、模倣学習時に模倣者が注目する行為の側面によって、(1) 型真似, (2) 行動単位模倣, (3) 目的行為模倣に分類しているが、各々において着眼して学習すべき事項の選択については未解決の問題として残されていることが指摘されている。しかし、正しい物理制約を推定する視覚機能と、その制約を達成させる動作プリミティブを組み合わせることで、予め知識として有している物体操作に限られるものの、適応的な学習が可能なが実証的に示されている (文献 [68] 等)。

これらのうち物体操作の視覚教示あるいは模倣研究について、ロボットの獲得す

る情報に応じて分類すると、軌道レベルでの教示研究と作業目標レベルの教示研究が存在すると考えられる。ここで軌道レベルでの教示は、観測した対象の位置・速度の変化、あるいは対象の把持方法を主として学習・再生するものとする。このような教示はコーディングの難しい道具利用法、あるいは真に最適あるいは重要な対象の動かし方を人から学ぶための手法として研究されており、文献 [17, 50, 94, 133] 等が挙げられる。Miyamoto ら [50] は、剣玉動作をおこなう人間の腕の数力所を視覚追跡し、滑らかかつ提示動作が可能な経路点を決定していく手法を提案している。佐藤ら [133] は、腕に取り付けた磁気センサや机上の力センサを利用して、人の物体操作軌道を推定することにより、物体のピックアッププレスやスライド動作軌道を得て再現している。Calinon ら [16, 17] は、物体追跡結果と手先の追跡結果、あるいは直接教示時を組み合わせ、ロボットの関節角度に関して次元圧縮をおこない、バケツ運びやチェスの駒の移動等における、真に最適あるいは必要な軌道を得た。岡田ら [94] は、移動プランニングを含む高次模倣行動システムを提案している。特に物体操作に関しては、認識された物体位置と、手を引いて覚えさせた手先位置により、物体の把持方法を記憶させている。

一方で、作業目標レベルの教示研究では、ロボット自身が有している物体認識機構や動作単位を当てはめて、比較的複雑な作業を再現させるための、作業の手順やゴールを教える用途とする取り組みがおこなわれている。文献 [34, 44, 108, 68] 等が挙げられる。これらの研究では、対象の近接・接触関係を用いてタスクの分節化をおこない、対応するロボットの動作を適用する枠組みとなっている。Kuniyoshi ら [113, 44] は、ピックアッププレスに基づくブロックの移動や積み上げに関して、ステレオカメラで人間の手とブロックを実時間追跡しながら、提示動作を分節化し行為を複数の階層で認識する手法を提案し、提示動作の因果関係の推測結果から積み上げ動作を再現することに成功した。Ikeuchi ら [34] は、適切な接触関係を達成するという組立作業を取り上げ、カメラやレンジファインダで物体を追跡し、接触状態を推定した。推定された接触状態変化を達成する動作を用いて、ブロックの挿入動作に成功している。その後も、動作プリミティブあるいはサブスキルを用い、三次元物体追跡結果を用いて、接触状態変化を伴うブロックの操作をおこなっている [68, 118]。また、小川原ら [108] は、多眼マルチベースラインステレオ装置を用いた物体の三次元追跡結果と、視体積交差法を用いた、部分的な遮蔽に対応した手先の

姿勢認識結果を用い、複数回の提示動作を比較して必須相互作用を推定した。接触状態の推移に関する誤差の修正を組み合わせ、注ぎ動作を含む動作の再現に成功している。提示軌道を接触状態変化に基づき修正しているため、作業目標レベルの教示と言えるだろう。

生活環境で働くロボットにとっては、双方ともに重要な機能だと言え、いずれも人間が人間に教え教わる方法と類似していることから、ユーザを選ばない教示方法だと考えられる。

しかしこれらの学習法を生活環境に適用する場合、観察対象を視覚認識可能であることを前提としていることが問題となる。特に軌道レベルの学習においては、対象の軌道そのものが得られなければ学習をおこなうことはできない。しかし前述の通り、日常環境内の道具には、上下・包含・突入関係を変化させるものが多く存在するため、対象に遮蔽が起こる状況は考慮すべきだと考えられる。

一方で、物体が物体へ切り込んでいくような関係の推定にも課題が残る。従来の視覚学習法における複数対象同士の関係推定は、剛体同士の位置と位置、あるいは面と面の関係に限られていると言える。日用道具の例として、包丁などは対象に切り込むことができ、切り込んでいること、そしてその軌道から得られる切断面同士の関係が、タスク再現においても重要な不変事項となる。さらに、布袋のような柔軟物も多く存在するが、変形するような対象では、完全なる三次元形状を取得することは難しい。詳細な三次元形状に頼ること無く、他の対象が中へ入っているか等の、高次の推定をおこなう手法が必要となる。

このように、従来の観察学習を日常環境タスクへ導入する際の問題点としては、以下の二点が考えられる。

道具と操作対象の遮蔽関係について

これまでに述べた人の作業の視覚学習手法、あるいは人の行動観察手法では、ほとんどの場合、毎時刻きちんと対象物の軌道が獲得できていることが前提となっている。つまり、作業に出てくる対象物の視覚認識に常に成功していることを期待している。しかし実際には、多くの場合において道具と操作対象は近接あるいは接触して作用する。このような状況では、一方が一方を遮蔽するような関係が頻発するため、常に視覚認識が可能であることを前提とするのは難しい。

道具を利用することで発生する，観測しにくい上下・包含・突入関係関係変化

道具を用いて作業をすることで，物体を持つ，動かす，置くといったピックアッププレース以外の操作対象の変化が存在する．特に，対象の変形を含む作用を表す，突入関係の変化推定法には課題がある．文献 [5] や [24]，[131] では切断のような作用が推定できているが，道具と操作対象の軌道が得られておらず，動作生成に必要なパラメタ，つまり切断面のような情報を推定できない．また，三次元の位置姿勢を前提として複数対象間の関係を推定する手法では，切断時に限らず，布袋のように完全な形状を取得し難い柔軟物について，他の対象との関係性を推定できない．上下・包含・突入関係の変化を推定するための視覚機能と，それを利用した視覚学習法の構築が必要と言える．

一方でこの二点は，ロボットが道具を用いた動作をおこなう際にも同様に課題となる．この点は次節で詳しく述べる．

最後に，ロボットの行動・動作記述を多くの環境や目的，ロボット身体へ再利用可能とするための多階層記述の観点から，上記の視覚学習法について考察する．家事支援行動のためのヒューマノイドの記述レベルに関しては文献 [122] にて，文献 [138] のロボット言語記述レベルの分類法をヒューマノイドのために改良し，高次の階層より 1) 自律行動レベル，2) 作業目標レベル，3) 対象目標レベル，4) 身体制御レベル，そして 5) 関節制御レベルの五つに分割することが提案されている．このうち，自律行動レベルはロボット自身がタスクの目的を設定し実行していくレベルであり，観察学習システムにおいては観察学習とタスク再現そのものとなる．図 2.2 に，作業目標レベル以下の構造を示す．作業目標レベルはタスク開始時の環境状態に関して一般化するレベルであり，操作の因果関係に関する知識を有し，現在の操作対象の状態を含まない作業目標を入力とし，環境記号認識結果から作業の手順を決定していくレベルである．従って，前述の視覚学習のうち作業目標レベルでの学習をおこなっていくべきレベルである．対象目標レベルは，操作対象の状態に関して一般化するレベルであり，道具の把持点に関する情報や軌道に関する知識を有し，細かな対象の位置を入力として，ロボットがとるべき手先の位置姿勢等を決定していくレベルである．従って，前述の視覚学習のうち軌道レベルでの学習をおこなっていくべきレベルと言える．身体制御レベルはロボットの身体に関して一般化するレベルであり，ロボットの身体に関する知識を有し，手先位置などを入力として逆運動

学演算をし、関節角度を決定していく。関節制御レベルでは、各ロボットのデバイスに合わせ関節角度変化時の補間等をおこない、環境へ何らかの作用を及ぼす。

つまり、前述の視覚学習のレベルについては、どのレベルで視覚学習すべきか、という問題ではない。各レベルでの視覚学習が必要となり、具体的には図 2.2 の赤枠で示したような事項が学習すべき事項である。青枠で示した事項は、タスク再現時に確認し利用していくべき事項である。

本研究におけるシステムでもこのような多階層記述を踏襲するものとし、上下・包含・突入関係を推定できる視覚機能がどのように有用となるかを、作業目標レベル・対象目標レベル双方のタスク学習と動作学習にて検証していくが、この点については後述する。

2.3 ロボットの道具を用いた作業における問題点

一方で、生活支援ロボット [3, 9, 23, 56, 71, 80] が予めインプットされたタスク、あるいは観察学習したタスクを遂行するためには、家庭内にある人間のために作られた道具を使いこなしていく能力が求められる。しかし、従来の生活支援ロボットのタスク例は、多くの研究は物体のピックアップブレースに代表されるように、単一物体を把持し、直接位置を変化させる操作に基づくものが多い。本稿で着目するのは、道具で操作対象に影響を与えるような間接的な対象操作、あるいは複数の対象を操作し、それらの関係が重要となるような場合である。このような操作は、生活支援タスクで頻出するものである。

図 2.3 に、このような支援タスクの例を示す。図中 A) はオフィスでの封筒への書類の出し入れ、B) はキッチンでのフライ返しを使った目玉焼きのよそいつけ、C) は要介護者へのズボン着脱支援である。これらの操作は、対象同士の上下・包含関係を変化させる操作であると言え、これらの操作においては、図のキャプションに記述したような問題点が存在する。従って各々のタスクは、以下の状態への操作と成功・失敗の確認、失敗した場合のやり直しを含む学習が重要なタスクであると言える：

- A) 封筒の表・裏面の間に書類が挟まっている状態
- B) フライ返しの上に目玉焼きが載っている状態

C) 衣類の布の間に足先が入っている状態

対象同士の接触状態が重要となるマニピュレーションとしては、組立作業が挙げられる。先行研究 [67] では、peg-into-hole 操作において剛体の面同士の接触状態を監視し遷移させる手法について検討がなされている。このような組立作業の研究ではある特定の接触状態を維持しながら別の接触状態へ確実に遷移させる操作に重点がおかれている。一方で前述のように生活支援タスクでは、薄い対象を直接対象に挿入するようなタスク、動的な操作や柔軟性のある対象の操作など、失敗が起こる可能性が必ず存在するために失敗復帰動作や学習が必要になる操作が多く存在する。

基本的な重なり合う物体マニピュレーション研究としては、Graspless manipulation の一例として、アルミボードの上の物体を plastic sheet の上に載せる “Scooping up” 動作を実現した研究 [1] がある。一方、文献 [82] では、爪を持つロボットアームを開発し、机から紙やプラスチックカードを持ち上げる動作に成功している。これらでは専用の sheet や爪を使っており、動作の失敗の恐れがないためセンサによる確認は行われていない。人間と同じ道具を使っている例としては、トングでパンを持ち上げている例 [39] があるが、認識についての記述は無い。文献 [75] ではアイスクリームの scooping 動作を実現しているが、アイスクリームの容器は動かないため、認識は初めのアイスクリーム上面の確認に限られ、一定の力ですくうための力制御に重みがおかれている。文献 [9] では、ロボットがパンケーキを焼くための動作の一つとして、spatula をパンケーキの下に差し込んでひっくり返す動作を実現させている。spatula の押し付け力に注意が払われており、spatula がパンケーキの下にきちんと入ったか否かの確認は行われていない。

扱う対象の形状や摩擦が一定である場合には、操作中の道具や操作対象同士の関係推定はそれほど重要でない。しかし一般的な道具・対象の操作においてはこの限りではなく、失敗・成功の認識と失敗復帰あるいは動作学習が必要となるであろう。ロボットの生活支援タスクにおいて、このような失敗復帰や動作学習をおこなっている研究としては、環境構造の扱いにおける失敗復帰 [93]、ピックアンドプレース [70] における失敗予測・復帰、衣類操作に関する失敗復帰 [79]、T シャツの操作のための動作学習研究 [69] 等がある。しかし、各々の対象・タスク依存のモデルが使われているか、単一の対象位置を確認しているにとどまり、複数の操作対象間に一般的に有用な関係推定法については述べられていない。

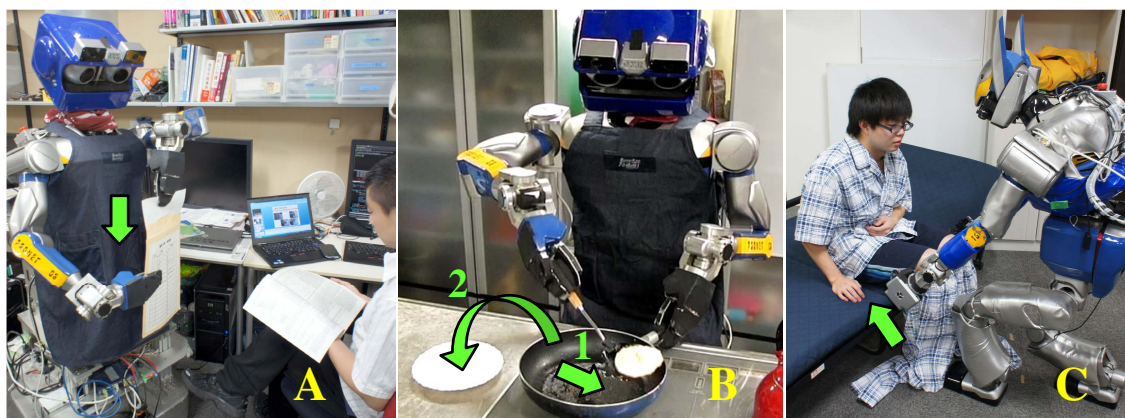


図 2.3: 生活支援ロボットによる日常タスクのうち、複数対象の接触関係の変化が重要となるものの例。それぞれ、A) 封筒へ書類を入れる、B) フライ返しを使って料理をよそいつける、C) 衣類の着脱を支援しているタスクである。それぞれのタスクには、以下のような問題点が存在する：

A) 封筒へ書類を入れる際、開口部が狭いので、認識・操作の誤差や紙の変形などで入らない可能性がある。書類が封筒の表裏面の間に入っていることを確認してから押し込む必要がある。

B) 目玉焼きの下へフライ返しを差し込む操作は静的な操作でないため、目玉焼きが動いてしまって載らない可能性がある。差し込む動作を行った後、持ち上げる前に目玉焼きがきちんと載っているか確認する必要がある。

C) 被介護者の足先にズボンや靴下を通す際、衣類の開口部は狭いので足が通らずに前後にずれる可能性がある。衣類を引き上げる前に、衣類の表面と裏面の間に足が通っていることを確認する必要がある（著者論文 [53] より）

ここで、一般に対象同士の下上・包含・突入関係を変化させる操作について、接触状態の変化に着眼した分類を行い、各々の操作における問題点を考察する。本項で述べる操作においては、ロボットが主として動かす操作対象を A とし、 A と近接・接触する操作対象を B とする。さらに、操作する前に B が空中に浮いている場合を除くと B は A に接触する前には、何らかの力（重力、クーロン力等）によってもう一つの物体または環境に接触していると考えられるから、これを C とする。これらの物体の間で重要な接触・近接関係がある場合とそうでない場合に分類する。ここで A と B に重要な接触・近接関係が存在するとは、 A が重力に反して B を支えて

いる状態， A の間に B が挟まっている状態， B が A の中に納まっている状態等，その関係なくして操作が成功し得ない，対象同士の重なり関係が存在することを指すものとする．

ここで新しい記号 Δ および \odot を導入する．ある面 a ともう一つの面 b に重要な接触・近接関係が存在する場合には $a \Delta b$ ，そうでない場合には $a \odot b$ と表すものとする．

以下，道具を用いた対象の操作，あるいは複数対象の操作を，操作で重要となる接触関係の推移に基づいて Op. 1 から Op. 6 までに分類する．さらに各状態変化での操作の問題点とその解決方法について述べる．ただし， a, a_i, a_j は物体 A の面の一部， b, b_i, b_j は物体 B の面の一部，同様に c は物体 C の面の一部とする．また，図 2.4 はこれらに対応する状態の例を示す．

Op. 1: $(A \odot B) \wedge (b \Delta c) \Rightarrow (a \Delta b) \wedge (b \Delta c)$

A が B に拘束されるように A を置く． A が正しく B の上に載ったり，かぶさったりするようにする必要がある． A や B が小さかったり， A と B の大きさが近いときにズレて失敗する．このときはズレに応じて A の位置を変える必要がある．

観測する際には， A に遮蔽されて B が見えなくなる，あるいははじめに B が見えないということが頻発するが， A と B の位置合わせには，その位置姿勢の特定が必要となる．

Op. 2: $(a \Delta b) \wedge (b \Delta c) \Rightarrow (A \odot B) \wedge (b \Delta c)$

B に拘束されていた A を持ち上げる．Op. 1 の逆の操作である．

観測の際には Op. 1 と同じ問題が発生する．

Op. 3: $(b \Delta a) \wedge (B \odot C) \Rightarrow (b \Delta a) \wedge (b \Delta c)$

B が C にも拘束されるように， B を拘束した A を動かす． B が正しく C の上に載ったり，入ったりする必要がある． B や C が小さかったり B と C の大きさが近いときにズレて失敗する．このときはズレに応じて A の位置を変える必要がある．

観測する際には， C は A と B に， A は B に遮蔽されて見えない状況が頻発する．しかし A, B と C の位置合わせには，その位置姿勢の特定が必要となる．

Op. 4: $(b \Delta a) \wedge (b \Delta c) \Rightarrow (b \Delta a) \wedge (B \odot C)$

C にも拘束されていた B を， A だけで拘束する．Op. 5 の後に起きる，Op. 2 の逆の操作である． A できちんと拘束できていなかった場合には落とすなど失敗する

ので、その場合は Op. 5 からやり直す必要がある。

観測の際には Op. 3 と同じ問題が発生する。

Op. 5: $(A \otimes B) \wedge (b \triangle c) \Rightarrow (b \triangle a) \wedge (b \triangle c)$

C のみに拘束されていた B を、 A でも拘束する。 A が小さかったり、摩擦力の読みが甘かったりすると正しく A で拘束することができず、 B がズレたりする。この場合は操作をやり直す必要がある。

観測する際には、 C は A と B に、 A は B に遮蔽されて見えない状況が頻発する。しかし A 、 B と C の位置合わせには、その位置姿勢の特定が必要となる。

Op. 6: $(b \triangle a) \wedge (b \triangle c) \Rightarrow (A \otimes B) \wedge (b \triangle c)$

B から A を離して、 C のみに拘束させる。Op. 5 の逆の操作である。摩擦力によって B が A にくっついてきたりすると正しく拘束を解くことができない。この場合は操作をやり直す必要がある。

さらに、操作前の B の拘束状態に関わらず、 A と B の複数の面が接触・近接する関係となる操作には、突入・包含関係を変化させる以下の二つがある。ただし、 a_* と b_* はそれぞれ、 A と B の面の一部とする。

観測の際には Op. 5 と同じ問題が発生する。

Op. 7: $(A \otimes B) \Rightarrow (a_i \triangle b_i) \wedge (a_j \triangle b_j)$

B が袋状になっている場合は入れる操作、なっていない場合は新しい接触面を作り出す、切るや刺すなどの突入関係を変化させる操作となる。袋状の B の面が薄い場合、あるいは B の表面に近い部分に A を差し込む場合には、認識や動作のずれにより A が正しく B の中に入らず、前後にずれて失敗する。このときは、前後の位置を正しくして A を操作しなおす必要がある。

観測の際には、 A が B に遮蔽されるが、 A と B の正確な位置合わせと接触関係の正しさを吟味するために、位置姿勢の特定が必要となる。

Op. 8: $(A \otimes B) \Rightarrow (b_i \triangle a_i) \wedge (b_j \triangle a_j)$

それまでの B の拘束如何に関わらず、 A を B の外側に被せる、あるいは被せて支える。 A の面が薄い場合など、認識や動作のずれにより B が正しく A の前後へと送り込まれず前後にずれて失敗する。このときは、前後の位置を正しくして A を操作しなおす必要がある。

観測の際には、 B が A に遮蔽されるが、 A と B の正確な位置合わせと接触関係の

正しさを吟味するために、位置姿勢の特定が必要となる。

このように、「入れる」「載せる」「切る」に代表される、複数対象間の上下・包含・突入関係の変化が伴うような、接触関係の変化をおこす操作においては、対象物の柔軟性や対象同士の摩擦等により、単一の動作のみでは失敗が起こり得る。動作の成否を認識する機構が必要であるが、対象同士に大きく遮蔽関係が発生していることが多い。その場合に操作対象の位置姿勢の特定は難しいため、遮蔽下でも操作対象を認識する、あるいは操作対象同士の関係を推定するための視覚機能が必要であると考えられる。

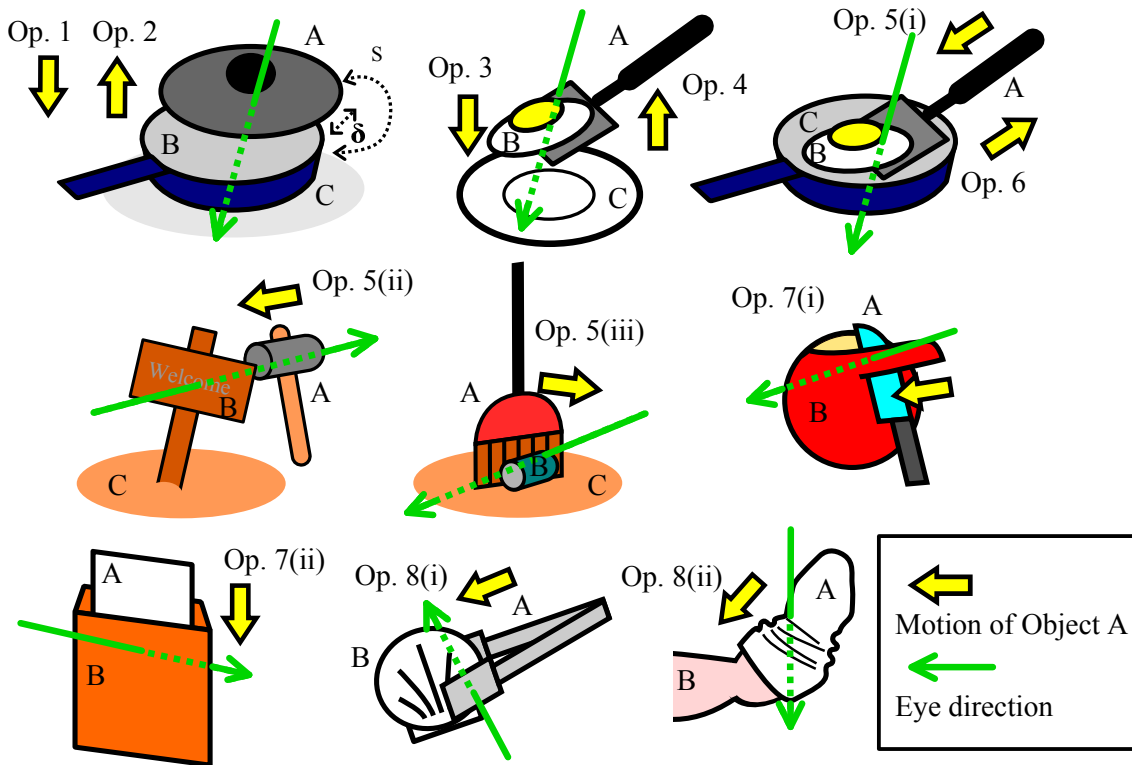


図 2.4: 道具と操作対象，あるいは複数対象間の重要な上下・包含・突入関係の変化，それぞれの例．

Op. 1: 蓋^Aをフライパン^Bの上に被せる．

Op. 2: 蓋^Aをフライパン^Bから取る．

Op. 3: フライ返し^Aに載せた目玉焼き^Bを皿^Cの上へ持って行く．

Op. 4: フライ返し^Aに載せた目玉焼き^Bを皿^Cの上から持ち上げる．

Op. 5: (i) フライパン^Cの上の目玉焼き^Bをフライ返し^Aですくい上げる；

(ii) 地面^Cにささっている標識^Bをハンマー^Aで叩いて真っ直ぐにする；

(iii) 床^C上のゴミ^Bを箒^Aで掃く．

Op. 6: フライ返し^Aの上の目玉焼き^Bをフライパン^Cの上にすべり落とす．

Op. 7: (i) リンゴ^Bの皮を包丁^Aで剥く；

(ii) 書類^Aを封筒^Bの中へ入れる

Op. 8: (i) トング^Aを使って貝^Bを持ち上げる；

(ii) 被介護者の脚^Bへ靴下^Aを履かせる．

2.4 視覚重畳関係を利用した作用推定

2.4.1 作用推定と対象の遮蔽

前節までで指摘したように，家庭環境内での道具利用では，対象同士に重要な上下・包含・突入関係の変化が多く起こる．しかし，観察学習時の観察対象あるいは操作時の操作対象が他の何かに遮蔽されている場合，それらの位置姿勢の特定は難しくなる．一方で，袋にしまう等，遮蔽関係を発生させること自体が重要なタスクも多くあることから，物体が物体に隠されている状況においても，それらの物体同士の関係を適切に推定することが重要となる．

しかし，このような状況において，カメラに代表されるような，対象からの反射波をセンシングする機器（反射赤外光を利用する RGBD カメラ等を含み，以降一般的なカメラと称する）では，直接に空間関係を把握するのは不可能である．電子タグを対象物に設置するなど，他の物体とセンサの間に遮蔽物があっても位置を推定できるような特別な枠組み [126, 129, 134] を利用することは根本的な解決策の一つであるが，道具だけでなく，家庭内の操作対象物一般（野菜，ゴミなどを含む）にもタグをつけていくことは難しいと考えられる．

複数のカメラを利用することで死角を減らす試みも数多くおこなわれており，室内での人物の動作観測等で古くから用いられている [128]．複数カメラシステムにおいては，相互オクルージョン関係にある剛体で交わりのない領域同士の位置特定に関する解析と立式化もおこなわれている [130]．しかし，前述のような袋に対象が完全に入り込む場合や，包丁が野菜に切り込んでいくような状況であれば，前者においては一方の対象が完全に，後者においても一方の対象の一部が，いずれの視点からでも観測不可能となってしまう．

上下・包含・突入関係を認識する手法としては，距離カメラや視差を用いる手法も考えられる．廉価な Microsoft Kinect [49] 等も入手可能になっているが，衣類や封筒の表面などの薄い対象同士の上下関係を安定に認識できるほど精度が良いわけではない．また，上下関係に起因する遮蔽は必ず発生し，視点の向こう側に存在する対象が直接観測できないことは通常のカメラを使う場合と変わらないため，何らかの手がかりによって見えない対象が存在することを推定する枠組みは別に必要となる．視差を用いる手法も同様である．

ロボットが自らの道具操作中に，上下・包含・突入関係を確認する手法としては，力センサ等による推定手法も考えられる．例えば対象Aが対象Bの中に入っているか否か，Aを上下に動かし，力がかかるか否か確認することで関係を推測するものである．このような方法は，生活支援ロボットにおいても，野菜が本当に切れているかを包丁を横に動かしながら力センサを使って確認する場合に用いられている [81]．また，柔軟物を一枚だけ掴んだことを strain gauge sensor によって確認している例として，[58] が挙げられる．しかし，操作対象と操作対象の接点がロボットの力センサから遠い場合，操作対象が軽い場合や柔軟性を有する場合には十分なセンサ値の変化は望めない．また，この手法は観察学習の際には用いることができない．

一方で本研究では，このように観測点と観測対象の間に他の対象が入った場合に，観測対象が遮蔽されて観測されないという関係自体が，観測対象と他の対象の関係を推定するのに有用であると考えられる．例として，図 2.5 のような，青い布と赤いペンギン（ぬいぐるみ）の関係を考える．(A) の右側の三つの段は，それぞれ青い布とペンギンの関係を変化させたものである．人間が見ると，一段目ではペンギンが青い布の上に乗っていることが分かる．一方で，二段目と三段目ではペンギンが見えないことは分かるものの，布とペンギンの関係は不明だと言える．(B) は，各段の右側の状態に移る直前の状態を撮影し追加したものである．これによると，二段目では人間がペンギンを青い布の下に入れていることが分かる．さらに三段目では，実は青い布が袋であって，ペンギンはその中へ入れられたということが分かる．この三つの事象を，物体の遮蔽関係の観点から考察する．すると，一段目の「ペンギンが青い布の上に乗った」状態は，ペンギンが青い布の一部を隠していることから推定できる．二段目の「ペンギンが青い布の下に入った」状態は，その過程でペンギンが青い布に隠されていることから推定できたものと考えられる．そして三段目の「ペンギンは青い袋の中へ入った」状態は，その過程でペンギンが布を隠しながら布もペンギンを隠すような状態を通り，最終的にペンギンが布に隠れたことで推定できたのではないだろうか．さらに「青い布が袋であった」ことは，ペンギンが青い布の中へ入れることから推定できたものとも考えられる．

以上のことから，物体が遮蔽されていること，そして物体同士の遮蔽関係とその時系列変化は，それ自身が物体同士の関係を推定するための重要な手がかりとなっているのではないかと考えた．

しかし、これまでも述べた通り、観測対象の領域同士が重なり合っている場合には視点から遠い対象は遮蔽下にあり、対象同士の重なり関係を、それらの位置や姿勢を直接観測した結果に基づき計算することは難しい。そこで次に、カメラの画像を入力として、対象領域同士の重なり関係、つまり視覚重畳関係を推定し用いる手法について検証する。これは画像から対象の表面同士の相対的な奥行き関係を推定するということである。画像から奥行きを認識するための手がかりとしては、文献[40]で以下のものが指摘されている：a) texture gradient, b) size perspective, c) motion parallax, d) aerial perspective, e) surface shading variations, f) relative upward locations in the visual field, g) outline continuity. このうち、複数の対象の距離関係を扱う場合は、a) は適さない。また、遮蔽関係が発生している場合には b) はそのまま適用できない。また、対象同士の距離が近い場合には c), d), そして e) も適さない。したがって、f) あるいは g) を用いるのが良いと考えられる。f) を用いている例としては、[40]で遮蔽関係から対象同士の重畳関係を推定する手法として Rosenberg et al. が提案した手法[60]を説明している。この手法ではまず画像内で領域分割した後、ある領域の輪郭が、その他の領域について何パーセント隠されているかに基づいて領域の順序を定めていき、“in-front-of”, “behind”, “equidistant” の関係を予測する。g) を用いる例としては、ある対象がある対象の境界線を遮蔽する際に出現する“T-junction”を用いて順序を推定する手法[59]が存在する。これらの手法には、既に複数の対象同士に遮蔽関係が発生している単一画像から、事前知識無しに対象の順序を推定できる利点がある。しかし低次元の視覚特徴のみから推定するために、推定精度には課題がある。一方で本研究の課題は、物体追跡中に、あるいはロボットが自らの手に持っている物体を操作中に、対象同士の上下・包含・突入関係を推定することである。この場合、各々の対象の全貌を予め見て記憶してから操作することが可能であると考えられる。そこで本研究では、f) を高次の特徴から推定する手法、つまり一度全体を見たことのある対象について、見えていたときの視覚特徴を用いて、遮蔽下において最も視覚画像を良く説明できるような対象同士の順序を推定していく手法を用いるものとした。

2.4.2 視覚重畳関係とその種類

本稿において（視覚）重畳関係とは，ある観測点から観測された複数の物体同士がその投影面において，互いにそれらの領域を遮蔽する，または遮蔽される関係を表すものとする．遮蔽関係は，ある物体の観測点側の表面が，もう一方の透明でない物体の観測点側の表面より遠方に存在するときに発生する．従って重畳関係は，1) 物体同士に遮蔽関係が存在する場合，または 2) 物体が物体に突入している場合，に発生する．

この重畳関係を，投影面における二次元の領域の重なり方によって分類すると，以下の三種類となる．

独立/Independent，または接触/Contiguous

物体領域同士に重なり合いが無い場合．

片側遮蔽/Covered

ある物体領域が，他の物体領域を遮蔽している場合．

相互重畳/Mutual

ある物体領域が他の物体領域の一部を遮蔽し，逆に他の物体領域から一部の領域を遮蔽されている場合．

これらの例を，図 2.6 に示す．図中 3. では，包丁が対象の一部を，対象が包丁の一部を隠しているので，相互重畳となっている．

一方，物体同士が三次元空間上でどのような関係にある場合に，どのような重畳関係が観測され得るかを示したのが，表 2.1 である．表では，凸かつ平面状の道具と凸形状の対象の三次元的な関係が，独立（Independent），接触（Contiguous），道具の一部が操作対象に突入している場合（Partially inserted），道具が全て操作対象に没入している場合（Fully inserted）の四種類を取り上げた．ただし，道具を A ，対象を B としたとき，各々の定義は以下のである．道具 A は凸形状の平面であるとし，

三次元 Independent 全ての道具上の点 A_i が，対象 B 領域外部に存在する．

表 2.1: 道具と操作対象が様々な三次元位置関係にある状況での、観察され得る視覚重畳関係．特に道具が凸形状かつ平面，操作対象が凸形状であった場合について示す（著者論文 [53] より）

		Tool		Object		Invisible edge of Tool		Region of Tool inside Object	
State of Superposition	Independent	Fully visible	Fully visible	Fully visible	Fully visible	Fully visible	Fully visible	Fully visible	Fully visible
One-way	Partially visible	Fully visible	Fully visible	Fully visible	Fully visible	Fully visible	Fully visible	Fully visible	
									Mutual
Operational Object	Fully visible	Fully visible	Fully visible	Fully invisible	Partially visible	Fully visible	Partially visible		
Tool	Fully visible	Fully visible	Fully invisible	Fully visible	Fully visible	Partially visible	Partially visible	Partially visible	
Contiguous									
Partially inserted									
Fully inserted									

三次元 Contiguous 全ての道具上の点 A_i は、対象 B 領域の外部もしくは対象 B の表面上に存在する。

三次元 Partially inserted 道具上の点 A_i について、対象 B 領域の内部に存在する点と対象 B 領域の外部に存在する点が双方存在する。

三次元 Fully inserted 全ての道具上の点 A_i は、対象 B 領域内部に存在する。

ただし、領域の内部と外部は、いずれも領域の表面を含まないものとする。従って、三次元 Contiguous 関係は上下関係を、三次元 Partially inserted は突入関係あるいは一部包含の関係を表す。また、三次元 Fully inserted は完全に包含される関係を表すものである。

各々の場合において、観測点を移動していったときに、どの重畳関係が観測され得るか、例を挙げながら示した。斜線部は、該当する重畳関係が観測され得ない場合である。ただし、観測点は道具平面外にあるものとする。このとき、相互重畳関係が観測され得るのは Partially inserted の場合に限られることに注意されたい。Partially inserted は、道具で操作対象を切断する、穴を開ける等の作用が行われる際に観測される。このような道具としてはナイフ、フォークや串などが挙げられるが、いずれも物体に突入していく道具の先端部は凸形状かつ平面状である。従って、物体を変形させる、特に切断させるような作用の認識には、相互重畳関係を観測することが有用であると考えられる。ただし、このような関係は道具と操作対象によらず、一般的な二つの観測対象同士について成り立つものである。

さらに、観測対象を凸平面と凸領域に限らず、二つの観測対象をいずれも任意の体積を有する形状とすることを考える。三次元 Independent、三次元 Contiguous、三次元 Partially inserted、および三次元 Fully inserted の定義は以下ようになる。二つの対象を A 、 B として、

三次元 Independent A 領域および B 領域に交わりがなく、表面上の点にも共通点が存在しない。つまり全ての A 領域内の点 A_i について、 $A_i \in B$ でない。さらに $A_i = B_j$ となる B 表面上の点 B_j は存在しない。

三次元 Contiguous A 領域および B 領域に交わりがないが、表面上に共通点が存在する。つまり全ての A 領域内の点 A_i について、 $A_i \in B$ でなく、 $A_j = B_j$ と

なるような A, B 表面上の点 A_j, B_j のペアが存在する。

三次元 Partially inserted A 領域および B 領域に交わりがある。つまり A 領域内の点であって、 $A_i \in B$ であるような A_i が存在する。しかし、全ての A 上の点について成り立つわけではない。

三次元 Fully inserted A 領域が B 領域に完全に含まれる、あるいは B 領域が A 領域に完全に含まれる。前者では A 領域内の全ての点 A_i について、 $A_i \in B$ が成り立つ。

この場合、(1) 三次元 Partially inserted のときの One-way (道具: Fully visible, 操作対象: Fully invisible) と One-way (道具: Fully invisible, 操作対象: Fully visible), (2) 三次元 Independent と三次元 Contiguous のときの Mutual (道具: Partially visible, 操作対象: Partially visible) は存在することとなる。その他の状態が存在しないことは、以下の 1. から 3. のように分かる。

1. 二つの体積を有する対象同士が Partially inserted か Fully inserted の場合に、二次元独立・接触関係が観測されないことは、証明 1 に記述するように証明できる。

証明 1

二つの体積を有する対象同士が Partially inserted か Fully inserted の場合に、二次元独立・接触関係が観測されると仮定すると、二次元観測画像上に曲線をひいて観測画像を二つの領域に分割したときに、各々の領域に観測対象が一つずつ入るようにできる。このとき、曲線の上に投影され得る点の集合と、画像の上下左右の枠に投影され得る点の集合によって、二つの領域が取り得る三次元空間も三分割（画角内部を C, D とし、外部を E とする）されており、観測対象の一方は C 、もう一方は D に含まれる。観測対象の一方を A 、他方を B としたとき観測対象 A 領域内の全ての点 A_i は、 B の内部に含まれ得ない。これは Partially inserted および Fully inserted の定義に反するため、二つの体積を有する対象同士が Partially inserted か Fully inserted の場合に、二次元独立・接触関係が観測されないことが証明された。

特に独立関係の場合には、曲線を交わらないよう二本引くことで領域 C と領域 D が接しないように領域 F をとることができるため、同様に A と B が三次元 Contiguous であることはない。

2. 体積を有する対象 A が体積を有する対象 B に完全に含まれる（三次元 Fully inserted）とき、二次元重畳関係は One-way (B が完全) に限られることは、証明 2 に

記述するように証明できる。

証明2

対象 A が対象 B に完全に含まれる場合，定義から A 表面上の全ての点 A_i について， $A_i \in B$ が成り立つ．観測点が B に含まれない空間上の任意の点 O とした場合， A_i と O を結ぶ線分は，必ず B 上の表面を通過する．従って A_i は O から直接観測され得ない．逆に B 上の任意の点 B_i について， B に含まれない空間上の任意の観測点 O と結んだ線分 B_iO について考えると， B_iO の間に A 上の点 A_k が含まれる可能性はあるが，やはり A_k と O を結ぶ線分は，必ず B 上の表面を通過する．従って， B が A に遮蔽されることはなく，体積を有する対象 A が体積を有する対象 B に完全に含まれる（三次元 Fully inserted）とき，二次元重畳関係は One-way（ B が完全）に限られる．

次に，相互重畳関係について考える．二つの対象が体積を有する場合には，双方の対象が三次元独立あるいは三次元接触関係にあった場合でも，二次元的には相互重畳関係が観測され得る．ただしこの場合には後述3. に示すように，少なくとも一方の対象が凹形状である．そしてその形状を含むような凸形状同士は三次元 Partially inserted 関係となる．相互重畳関係かつ重畳領域が十分な場合には，対象の一方には十分な凹みがあると考えられ，その凹みへもう一方の対象が入りかけている，あるいは三次元接触関係の場合には入っているということとなる．従って，二つの対象が体積を有する場合においても同様に，二次元の相互重畳関係を手がかりとして，物体同士の一部包含関係を推定しても問題ないと考えられる．

3. 二つの突入・包含関係の無い物体が二次元的に相互重畳関係にある場合に，少なくとも一方の物体が凹形状であることは，証明3に記述するように証明できる．

— 証明3 —

一方の物体を A ，もう一方の物体を B とする．二次元的に相互重畳関係が発生しているならば，投影された画像において，物体 A を射影した領域のうち物体 B の領域を遮蔽している領域と，物体 B を射影した領域のうち物体 A の領域を遮蔽している領域が存在する．前者の領域内部の点を I_1 とし，後者の領域内部の点を I_2 とする．また，カメラ中心を O とする．さらに I_1 に写る物体 A 上の点，つまり物体 A の手前側の点を A_1 とし， OA_1 を点 A_1 側に伸ばした直線と物体 A が再び交わる点，つまり物体 A の奥側の点を A'_1 とする． OA_1 をさらに点 A'_1 側に伸ばすと，物体 B の表面に交わるが，同様にその手前側の点を B_1 ，奥側の点を B'_1 とする．二つの物体に包含関係が無い場合には，包含関係の定義上， $O-A_1-A'_1-B_1-B'_1$ の並びとなる点 A'_1 ，点 B_1 および点 B'_1 が必ず存在する． I_2 についても同様に， $O-B_2-B'_2-A_2-A'_2$ の並びで一直線となる A_2 ， A'_2 ， B_2 ， B'_2 が必ず存在する．ここで，もし二つの物体が凸形状であったとするならば，各々の物体において，表面の任意の二点を結ぶ線分は必ず物体の内部のみを通る．そこで A_1 と A_2 ， A'_1 と A'_2 ，を線分でつなぐが， B_1 と B_2 ， B'_1 と B'_2 を線分でつなぐと，これは前者と交点を持ってしまう．これは，物体の内部のみを通るという事実に矛盾することから，二つの物体が凸物体であるという仮定が誤りであることが証明された．

ただし，上下・包含・突入関係は改めて以下のような関係であるとする．

上下関係

複数の物体同士が接触しているが，突入・包含関係は無く，片方の物体がもう一方の物体を支える関係にある場合を指すものとする．例えばフライ返しで目玉焼きを載せている場合，フライ返しと目玉焼きには上下関係が発生しているとする．

包含関係

単に包含関係と言った場合には，片方の物体領域に，もう一方の物体領域が完全に含まれる関係を指すものとする．一部包含の関係とは，片方の物体が凹形状をしており，その凸包ともう一方の物体の領域同士に交わりがあるものを指すものとする．例えば袋にぬいぐるみを入れた場合，袋とぬいぐるみの間には包含関係が発生しているとする．

突入関係

一方の物体が、もう一方の物体へ、その形状を微小変形させながら入り込んでいく関係のことを指すものとする。例えば包丁で野菜を切っているとき、突入関係が発生しているとする。

視覚重畳関係は、複数対象を観測した際に、あくまでも観測面での対象領域同士の間接関係を表すものである。二次元の関係であることから、観察対象の三次元的な状態を完全に推定するための手がかりとして十分ではない。しかし前述のように、少なくとも相互重畳関係を用いることで、一方の観測対象が他方の観測対象に突入する、あるいは一部が包含される状態を推定することができる。さらに、一般的な重力下では、一方の対象が一方の対象を支えている関係にある場合、それを人間の視点から観測した場合には、人間は手を目よりも下で動かすことが多いことから、それら二つの対象の間には重畳関係が発生している場合が多いと考えられる。これは等身大ヒューマノイドの場合でも同じであり、図2.4の緑線で表すような視線方向を用い、視覚重畳関係を観測し、さらに他の視覚的手がかりを併用し、関係推測のための必要条件を追加していくことで、観測対象同士の上下・包含関係を推定していくことが可能だと考えられる。具体的には、物体同士が近づいている場合に、下に存在する物体も一部は見えているような片側遮蔽関係が発生している視点が存在し、さらに物体同士に運動の随伴性が確認できれば、一方の物体が一方の物体を動かしている関係となっている可能性が高い。これはいずれも物体同士が接触し合い、片方が片方を支えている、という上下関係の推定のための必要条件だと考えられ、全てを確認することで、その信頼性が増すと考えられる。また、物体に物体が近づいて行った後に、片方が完全に見えない片側遮蔽関係が観測された場合、それは上下・包含関係が変化しているための必要条件となると考えられる。この関係の推測は、上下関係の下側の物体、あるいは包含関係の外側の物体が動かされたときに、上側あるいは内側の物体も随伴して動いていく関係によって信頼性を上げることができるであろう。

もちろんこれは全ての視線方向において観測され得るものではなく、例えば皿の上に料理を盛りつけたものを真横から観測する場合には、皿と料理の間には十分な重畳関係は観測されない。各々の視覚的手がかりは、観測対象の三次元的な関係を推定するのに十分ではないため、観測する視点によって、異なる関係が推定される

可能性もある。しかし家庭環境内で人間がおこなっている動作においては、その目的は三次元形状同士の詳細な相対位置関係を達成するより、人間の視点に基づいて観測され得る、視覚重畳関係の変化を達成することの方が重要である状況は多いのではないだろうか。例えば、物に布をかけるといった場合には、物と布の詳細な位置関係よりも、布が物を遮蔽していることの方が重要であろう。物をどこかへ片付けるという場合にも、物が外から見えなくなっていることが重要であることも多いと考えられる。ロボットにも人間と同様の視点を与え、このような関係を推定させることで、人間の感覚により合致したタスクの目的を推定できるようになるのではないかと考えた。

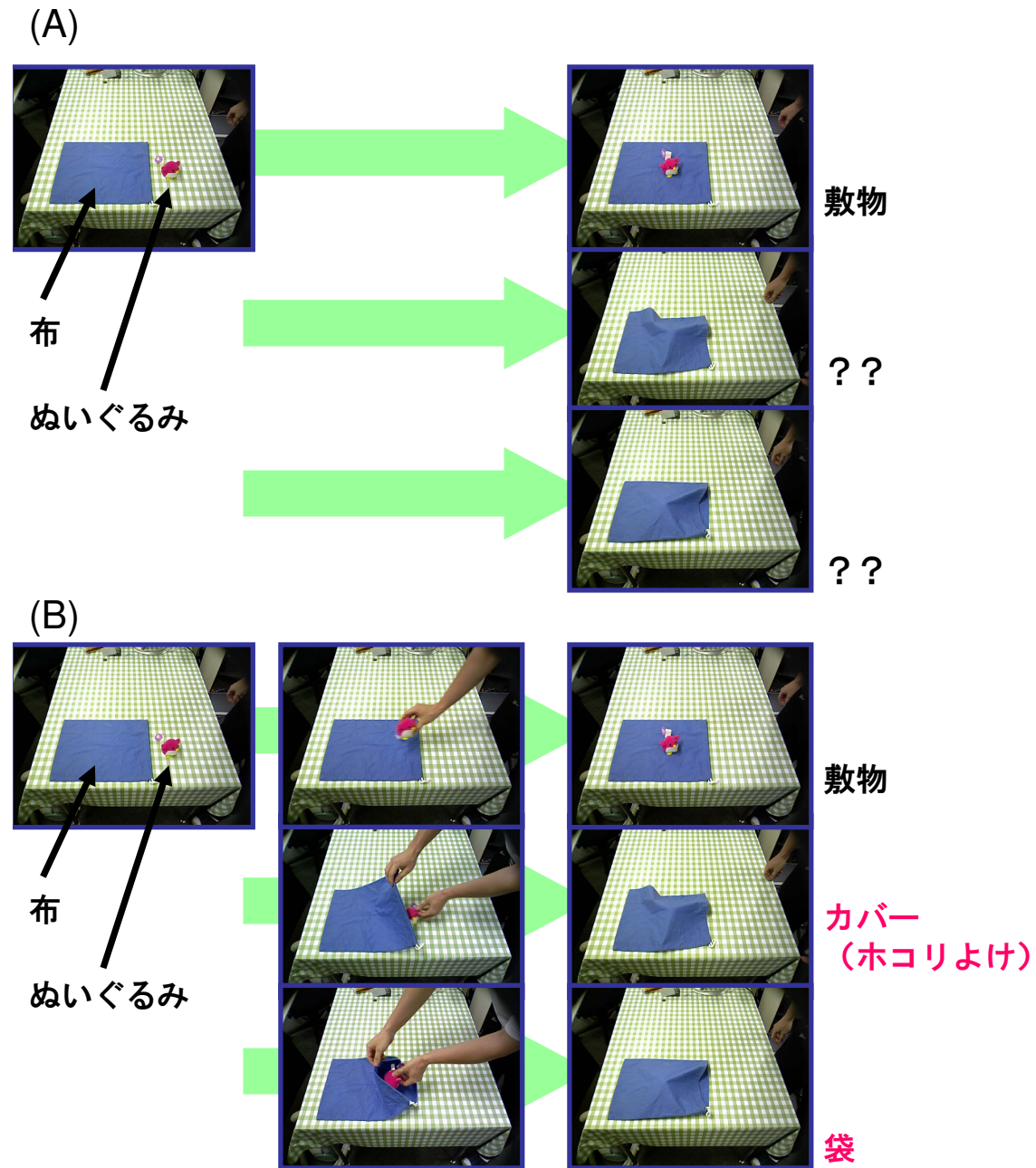


図 2.5: 遮蔽状況や遮蔽にいたる状況を用いた作用推定

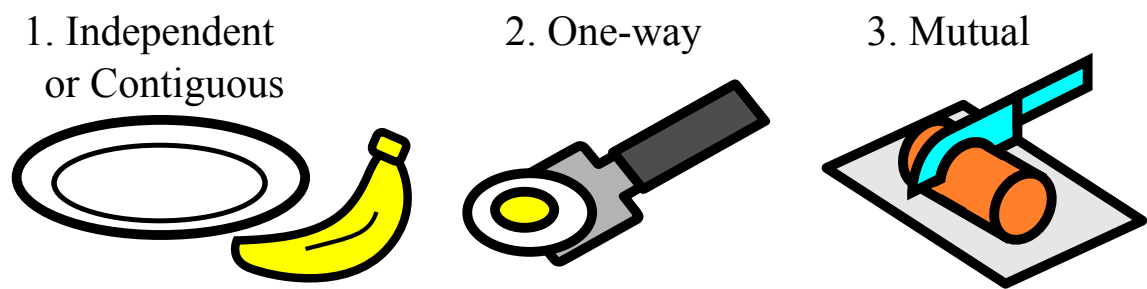


図 2.6: 視覚重畳関係の種類．それぞれ 1. 独立あるいは接触，2. 片側遮蔽，3. 相互重畳関係を示す

2.5 重畳関係を推定するための手がかり：空間遮蔽近傍情報と遮蔽時刻近傍情報

前節までの議論により、視覚重畳関係という視覚的手がかりを用いることができれば、観測している観察対象同士、あるいは操作対象同士の関係を推定するのに有用であると考えられる。しかし本章で述べてきた通り、観察対象同士あるいは操作対象同士が近接し合っている場合、視覚重畳関係やその他の手がかりを視覚により推定するためには、観測している対象に発生している遮蔽が問題となる。

ここで、一般的なカメラにて観測すべき対象を含む画像を撮影し、何らかの画像処理手法や追跡手法によって、観測対象上の点を特定することを考える。対象が観測できない場合、観測できる場合、あるいは認識できる場合、認識できない場合を分類し、その原因を考えると図 2.7 のようになる。

通常のタスク学習あるいは動作学習においては、正しい教師データを与えるために、図 2.7 で【VISIBLE+CORRECTLY RECOGNIZED】に分類される認識結果のみであることを期待する。一方で、これまでに論じた、道具を用いている状況下で起こり得る遮蔽というものは、上では【NON-VISIBLE-2】に分類され、対象上の点が画像へ写っていない状況であるため、そもそも画像内の視覚処理に基づく認識、追跡が難しいものである。しかし、観察対象の一部が【NON-VISIBLE-2】であっても、残りの一部が【VISIBLE+CORRECTLY RECOGNIZED】であった場合には、対象の形状の連続性を用いることで、後者の配置から前者の配置を推定することが可能ではないかと考える。ただし、この場合は【NON-VISIBLE-1】と【VISIBLE+UNRECOGNIZED】だけでなく【VISIBLE+WRONGLY RECOGNIZED1】、【VISIBLE+WRONGLY RECOGNIZED2】を十分に減らすことが必要となる。本研究では、対象の一部領域が遮蔽下にあっても、遮蔽されていない領域の特徴を用いて安定した対象追跡を可能とすることで、これを達成することを試みる。

一方で、観察対象が他の対象に完全に没入したときやフレームアウトした場合など、観測したい対象の全ての領域が観測できない状況では、このような手法も用いることができない。この場合においては、対象を観測可能であった瞬間の特徴のみから観測不能下での対象の状態推定をおこなう必要がある。

これらを整理し、本研究で提案して検証するのは以下の二種類の手法である。

空間遮蔽近傍情報の活用

遮蔽されている部分のまわりの空間の特徴を利用することで、観測できない対象の状態を推定する方法論である。例えば、キッチンでカレールーをお玉杓子でかき混ぜているのを見ているとき、お玉杓子の先の部分は、直接見ることはできない。しかし、お玉杓子の柄の部分は見えていることが分かり、なおかつそれを追跡し続けることができれば、柄からつながっている先の部分がどこにあるか、推測が可能であると考えられる。

この手法は、観察対象に見えている部分が存在するような重畳関係の場合に用いることが可能なものである。これはつまり、相互重畳関係と二次元片側遮蔽のうち双方に見えている部分が存在する場合である。この場合には、相互重畳関係が特定できれば、前述の通り重畳関係そのものから物体同士の突入関係あるいは部分的な包含関係が推定でき、観察者の視点を人間の視点と類似した高さ・位置であると仮定すれば、片側遮蔽関係より上下関係を推定できると考えられる。他に運動の随伴性という視覚的手がかりを併用することで、物体を支えながら動かす、物体に切り込むという関係の推定に有用であることを、本論文4章で詳しく述べる。

遮蔽時刻近傍情報の活用

観測している複数の対象が遮蔽に至るまでの時刻、あるいは遮蔽状態を解消する時刻の情報を利用することで、対象が観測されていない時刻の状態を推定する方法論である。例えば、ロボットが人を観察している際、人が持っているぬいぐるみがどこかへ動いている最中に見えなくなった場合を考える。この場合でも、見えなくなった先に布袋があれば、ぬいぐるみは布袋に収まったのだらうと推定できる。また、人が布袋を持って移動した際に、ぬいぐるみが布袋のあった場所に残されていないければ、ぬいぐるみが布袋と一緒に移動していったことが推定できる。

この手法は、観察対象に見えている部分が存在しないような重畳関係が発生する時刻がある場合に有用だと考えられる。これはつまり、二次元片側遮蔽のうち、片方の対象しか見えない場合である。この場合には、対象が対象に覆いかぶさっている状況や、対象が対象にすっぽりと入っている状況が含まれる。

このような対象の道具としての利用法は、片付けや物を運ぶ場合等で出現するが、対象が対象に入りあるいは載ることで、本当に一方の対象を運搬できるものであるかも、対象同士の運動の随伴性を推定し用いることが可能であれば、学習することが可能だと考えられる。本論文5章で詳しく述べる。

以降、これらを重畳関係、あるいは重なり領域を推定するための手がかりとして、手がかりを用いた重畳関係の推定法と利用法について実証的に明らかにする。

2.6 検証に用いるタスク学習システム

本研究では以上までの考察に基づき、日常タスクを学習しながらおこなっていくロボットが、物体の遮蔽関係という視覚的な手がかりを利用して、対象の状態推定をおこなうことの可能性と有用性について、実証的に明らかにする。本章での議論より、以下のような問題点をとりあげ実証することが、物体の遮蔽関係がタスク学習一般で有用であることを検証することに相当するものだと考える。

検証課題1：日常的な道具と操作対象の間の遮蔽関係の観測の可能性

前述の通り、視覚重畳関係は、複数の観察対象あるいは操作対象の関係性を推定するための視覚的な手がかりとして有益ではないかと考えられる。一方で重畳関係が発生していることは、一方の対象が遮蔽下にあることを表す。これは通常の視覚ではセンサに写りもしないという状況である。遮蔽されている対象あるいは対象領域がどこに存在するかを、見えている範囲内の情報のみを用いて推定するための手法が存在するか、あるいは存在するとしてどのようなものが問題となる。この手法が存在すれば、次項のように、重畳関係を対象同士の関係性変化の推定へ、視覚的な手がかりとして用いることが可能となる。

本稿では遮蔽下の状況を推定する際に、二通りの戦略を検証するものとする。それぞれ本章で述べた“空間遮蔽近傍情報”と“遮蔽時刻近傍情報”であるが、これらの詳細については4章と5章で述べる。さらに、このような関係を利用するための高視力視覚の構成法について、3章で述べる。

また、対象同士の重畳関係推定に先立って、対象の位置・姿勢や領域の特定が可能であれば、その追跡結果は、本章はじめで述べた軌道レベルでの道具操作

学習一般へ活用可能だと言える。この点についても4章で検証する。

検証課題2：視覚重畳関係を用いた物体同士の関係推定の可能性

前項で視覚重畳関係が観測できた場合に、その重畳関係を用いて、あるいはその他の視覚的手がかりとも組み合わせることで、物体同士の関係をどのように推定できるかについて議論する。このような推定が可能であり有用なものであれば、対象が遮蔽されているという情報そのものを用いることに有用性が存在するということとなる。つまり、本章はじめで述べた作業目標レベルでの分節化へ一般的に貢献するものだと言える。

本節後半で考察すると共に、4章と5章で実験をおこない検証する。

検証課題3：

視覚重畳関係を用いた物体同士の関係推定を用いた動作・タスク学習の可能性

前項で物体同士の関係推定をおこなえた場合に、それがどのようにロボットのタスク実行やタスク学習へ有効かを、実証的に議論する。特に4章では、失敗復帰を含む軌道レベルでの動作修正への活用を試みる。これは視覚重畳関係を用い、単一の操作に関する対象目標レベルの動作学習をしながら失敗復帰をしていくものである。一方で5と6章では、視覚重畳関係を基軸とした観察学習により獲得した作業目標レベルのタスク目的と物体の道具的利用法に基づき、実行時にロボットが自身に適した操作手順を決定しながら、一連のタスクを再現する実験をおこなう。いずれも、これまで学習が難しかった上下・包含・突入関係を用いた動作修正や、それらの観察学習および再現が可能となることが期待できる。

検証課題4：生活支援タスクの学習・実行に際しての有用性

前項までがおこなえた場合、そのようにロボットが遮蔽情報を用いることが、ロボットが生活支援タスクをおこなう上でどのように有用であるか、実タスクにより検証する。これには遮蔽情報の利点・欠点だけでなく、前述のタスク学習システムの特性と限界の評価も含まれる。

4章後半の実験では、料理のよそい付け作業とファイリング作業を、6章の実験では片付け作業と洗濯物の分類を扱う。

そこで本論文では、図 2.8 の上段に示すような道具利用タスク学習システムを用いて、上記課題を検証していくものとした。ただし、このシステムにおいて学習できる事項も限りあるものであり、他のレベルでの学習を否定するものではないことに注意されたい。道具と操作対象の間に発生する遮蔽関係についての知識を利用し、道具や操作対象の見えていない部分や見えていない時区間で発生している作用を推定し利用することが、従来よりも幅広いタスクや動作の学習を可能とするか否か、を検証するとともに、それ自身が有用な道具利用生活支援のタスク学習システムとなることを狙うものである。

図 2.8 で想定する観察学習の流れは、以下のものである：ロボットはまず、センサを頼りに観察対象の追跡をおこなう。追跡結果を用いて、対象間に現れる重要な変化（作用）を推定し、各対象に許される操作方法、そして最終的な対象の目標状態を推定し、記憶する。タスク再現時には、ロボットの体と再現時の環境から、学習した各対象に許される作用のさせ方（機能）を用い、学習した最終的な対象の目標状態を達成するための作業目標レベルでのタスクプランニングをおこなう。タスクプランナは予めロボットが持っている動作のセットから、適切な順序の動作列と操作対象を計算するものとし、ロボットは対象目標レベル以下の知識を用いて、計算結果の通りの動作をおこなう。ただし、ロボット自身の道具を用いた動作においても、次節で述べるように様々な問題点が存在し、失敗が起こり得る。そこでロボットは操作対象の状態を観測し、軌道レベルでの失敗復帰や学習をおこなう。これは前述の対象目標および身体制御レベルに相当する。最終的にロボット実機の動作時には、関節レベルの制御がおこなわれる。

このシステムでは、追跡結果から視覚重畳関係に基づき、対象間に現れる重要な変化を推定する部分が、視覚教示における作業目標レベルでの教示に相当する。一方で、対象追跡を連続的におこなうことができれば、それは視覚教示における軌道レベル（対象目標レベル）の教示の入力となり得る。また、ロボットが操作再現時に操作対象の状態を観測して失敗復帰をおこなう際も、成功・失敗状態推定に基づいてマニピュレータの軌道を変化させることは、軌道レベル（対象目標レベル）での学習となっていると言える。このように二つのレベルが組み込まれた学習システムを用いて視覚機能の評価をおこなうことは、軌道レベルと作業目標レベルの双方のレベルでの視覚機能の評価をおこなっていることに相当すると考えられる。従っ

て本研究で提案する視覚機能が，その中でどのように有用であるかを評価することは，視覚教示一般での有用性を評価していることに相当するものだと考えている．

各検証課題からこのシステムを見た際の各課題の範囲および本論文の構成は，図2.8の下段に示した．

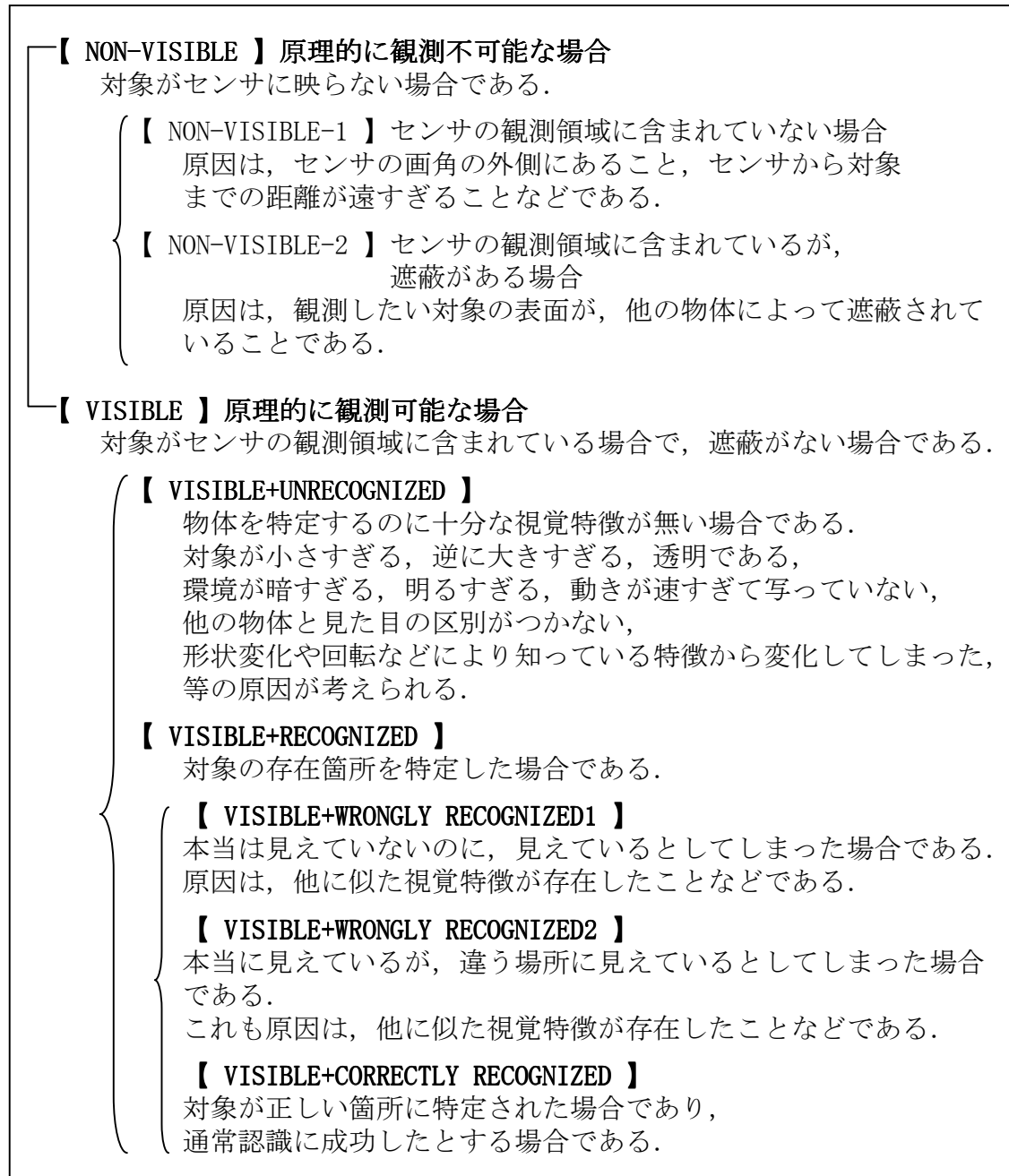


図 2.7: 一般的なカメラと画像処理手法によって観測対象上の点が観測できる場合、観測できない場合の原因別分類



2.7 道具利用タスク学習における用語

本論文では特別な断りのない場合，以下の用語を注釈のような意味で用いるものとする．

操作対象物，道具

観察学習時には人間が，ロボットによるタスク実行時にはロボットが，主として操作し，その操作がタスク達成に必要なような物体のことを，人間あるいはロボットの操作対象物と言う．操作対象物は複数存在することができ，その中でも他の対象に何らかの影響をおよぼすことが主な目的となる対象物を，道具と表現するものとする．特にロボットによる観察学習時には（ロボットにとっての）観察対象と表現するものとする．

作用

対象物や道具が人間から受ける，あるいは対象物が道具から受ける，タスク目的を達成するために不可欠な重要な影響，つまり状態変化のこと．例えば，入れられる・載せられる・切られる等．

道具的機能

物体が有するもので，他の物体へ影響を及ぼす（作用する）ことのできる能力のことを言うものとする．対象が何らかの道具的機能を持つときのみ，他の操作対象の状態を，その道具的機能の種類範囲内で，変えることができる．道具的利用法も同義である．このような操作のことを，その物体の道具的利用と呼ぶことにする．

動作・操作

操作対象に作用（状態変化）を及ぼすために必要である，作業目標レベル記述の最小のサイズのものを指すものとする．

タスク

操作対象を目的の状態に変化させるための，一連の状態変化のこと．観察時には，各観察対象に及ぼされる作用（状態変化）の時間的推移によって記述される．再現時には，状態変化を起こすための動作列によって記述される．

タスク目的（ゴール）

人間がロボットに学習させたい、作業の最終目標となる状態のこと。例えば物を運ぶタスクでは、そのタスク目的は「物が目的地にあること」となる。

失敗・失敗復帰

ロボットが一つの物体操作をおこなった後に、操作対象に目的の状態変化を及ぼせなかった場合を、失敗と呼ぶものとする。失敗の可能性の高い動作については、操作終了後に成功の可否を認識し、操作のやり直しを含む失敗復帰をおこなうべきである。視覚重畳関係を用いた失敗認識と失敗復帰に関して、4章の後半で詳しく述べる。

メタ情報

タスクの観察学習の際に、対象の認識結果から作用を推定するための中間データのこと。本論文4章と5章では、対象の認識結果からメタ情報あるいはメタ情報の時間的推移を計算し、各時刻のメタ情報あるいはメタ情報の時間的推移から、各時刻の作用を決定する。本論文で用いるメタ情報は、1) 認識の可否、2) 各観察対象の動きの有無、3) 対象同士の距離、4) 観察対象同士の運動の随伴性、5) 視覚重畳関係の5種類であり、1)、2) は対象ごと、3) から5) は対象ペアに一つ存在する状態で、各々時間毎に状態が存在するものとする。

2.8 おわりに

本章では、従来のロボットによる物体操作の視覚学習研究と生活支援ロボットによる道具操作についてまとめ、観察対象や操作対象が薄い、あるいは柔軟性を有する場合に、対象同士の上下・包含・突入関係の変化を、対象に遮蔽が発生している場合においても推定することの重要性と難しさを指摘した。そして、観察対象や操作対象同士の遮蔽関係である視覚重畳関係を陽に用いることが、タスク学習の際の対象の観測や対象同士の関係の推定に有用である可能性について考察した。次章以降では、以下のような4つの問題に整理してしていく。

1. 日常的な道具と操作対象の間の遮蔽関係の観測の可能性
2. 視覚重畳関係を用いた物体同士の関係推定の可能性

3. 視覚重畳関係を用いた物体同士の関係推定を用いた動作・タスク学習の可能性
4. 生活支援タスクの学習・実行に際しての有用性

第3章

道具を用いた生活支援タスクを学習する ロボットシステム

3.1 はじめに

本章では，日用道具を用いて生活支援を学習し実行するロボットへの要求機能を明らかにし，一つの実現例として，実験で用いる等身大ヒューマノイド“HRP-2VZ”の構成法を示す．さらに HRP-2VZ を用いた道具利用の基礎実験から，日用道具を用いて生活支援をおこなうロボットの問題点について考察する．また，広視野視覚と高視力視覚による協調注視制御実験では，ロボットが人の観察中に人の手元へ自律的に高密度の注意を向けることが可能かを検証する．

なお本章の記述は，著者論文 [51, 52, 149] の内容を含む．

3.2 日用道具を用いて生活支援をおこなうロボット

家庭内に入り，人の傍らで人の代わりに作業をおこなうロボットのハードウェアへ必要な要求機能として，大きく以下の二点が挙げられる．

日用道具を用いるためのマニピュレーション能力

人間が普段生活している環境では，キッチンやタンス，廊下や階段などの環境構造，そして包丁や雑巾などの道具や物品類は，基本的に人間が操作しやすいサイズや形状をなしている．このような環境内で，ロボットが人間と同等の物体操作をおこなうためには，ロボット自身も人間と類似した体の構造を有していることが望ましいと考えられる．具体的には，人間と同等の背丈を有すること，そして手先が三次元空間内で任意の位置・姿勢を取り得るようにするためには，マニピュレータは六以上の可動軸を有することが必要となる．

また，本論文6章で実験をおこなっているような，柔軟物を道具として用いるような作業においては，柔軟物の形状を保ちながら他の対象を操作するために，マニピュレータは双腕であることが望ましい．ただし，自由度の多い構造は高価でメンテナンス性が低くなる可能性が高い．単腕であっても，物を拾ってくる，扉を開ける等の簡単な介護作業の多くをおこなえるであろうことから，本論文6章では単腕ロボットでの物体操作についても検証実験をおこなう．さらに前章図2.2に示したように，そのソフトウェア構成は作業環境やロボット身体に適応できるよう，多層構造とするのが望ましいと考えられ，本システ

ムでも同様な構造とする。

人間並みの視力，三次元環境把握能力

人間の物体操作を観察するために，そして普段人間が利用している道具や物品類を人間と同等に操作するためには，物体位置の特定，形状認識や物体同士の関係の認識等に人間並みの視力を有することが望ましいと考えられる．特に本研究で活用する物体同士の視覚重畳関係は，人間が観測する際と同様の状態として観測することで，人間と同様の作用の推定へ有用だと考えられる．これには，視覚は人間と同等の箇所にも有することが望ましい．前項と合わせ，多自由度マニピュレータを有し，頭部に人間並みの視力，かつ三次元環境把握能力を有する視覚を備えることが望ましいこととなる．

人間並みの視力について考察すると，次章で扱う包丁と操作対象の重畳関係推定では 200[pixel]，ファイリング作業とよそいつけ作業における安定した重畳関係推定には，3000[pixel] の重畳領域が必要であった．手元で物体操作をおこなう通常の状態を仮定すると，例えば頭部カメラから 600[mm] ほど離れた場所において幅 10[mm] のものを 20[pixel] 程度で観測できるためには，約 $4.8 \times 10^{-2} [^\circ/\text{pixel}]$ 程度の角度分解能が必要であり，これは視力 0.35 に相当する．もう一つの例として，環境中の文字を検出し読むための視力について考える．画像内の閉輪郭領域として文字候補を検出して OCR で読み取りをおこなう手法 [78] を用いると，後述のズームヘッドの光学ズームカメラでは，文字の高さが 20[pixel] から 25[pixel] 程度で文字検出が可能であった．一般的なプリントの文字 (10.5pt) が手元にあると仮定し，カメラから 600[mm] の位置から検出できるようにするためには，約 $1.8 \times 10^{-2} [^\circ/\text{pixel}]$ の角度分解能が必要であり，これは視力 1.0 に相当する．

3.3 日用品を用いて生活支援をおこなうロボットの構成法

本節では，全節で指摘した物体マニピュレーション能力を有するロボットの例として，本研究で実験に使用した等身大ヒューマノイドロボット，“HRP-2VZ”につ

いて説明する．HRP-2VZの全身写真を図3.1に示す．

3.3.1 HRP-2VZのハードウェア構成

HRP-2VZは、ヒューマノイドプラットフォーム HRP-2W の下半身を全方向移動台車 V_{max} (株式会社小野電機製) に換装した HRP-2V[35] をさらに改造し、より広い視野と高い視力を実現するための頭部“ズームヘッド”を搭載した日常生活支援実験用ヒューマノイドである．したがって、頭部～上半身～下半身の順にズームヘッド～HRP-2W～ V_{max} という三つのシステムが存在している．

ベースとなった HRP-2W は、経済産業省の「人間協調・共存型ロボットシステムの研究開発」(HRP: Humanoid Robotics Project) [139, 140] の一環として開発された [32, 41]．東京大学大学院 情報理工学系研究科 情報システム工学研究室 (JSK) では、HRP-2W の両手首関節へピッチ軸を 1 軸追加し、腕を 7 自由度としている．また手首関節の可動範囲の拡大をおこなうことで、手先がとれる位置姿勢の範囲を拡大させ、遂行可能タスクを拡大している．これにより、HRP-2W と V_{max} を合わせた HRP-2V の関節構成は表 3.1 のようになっている．

表 3.1: HRP-2V の関節構成

部位	関節名称
胴体部	腰ピッチ軸, 腰ヨー軸
頭部	(胴体側より) 首ヨー軸, 首ピッチ軸
腕部	(胴体側より) 肩ピッチ軸, 肩ロール軸, 肩ヨー軸, 肘ピッチ軸, 手首ヨー軸, 手首ロール軸, 手首ピッチ軸, ハンド開閉軸 ×2
台車部	全方向移動台車 (omni wheel x 4)

3.3.2 ソフトウェア構成と道具利用のための動作生成法

ズームヘッドを搭載した場合の HRP-2VZ のソフトウェアシステムについて、図 3.3 に示す．ただし認識および動作生成に関しては例を挙げての簡略記述となっている．本研究では、前章図 2.2 で示したように、作業目標レベル記述と対象目標レ

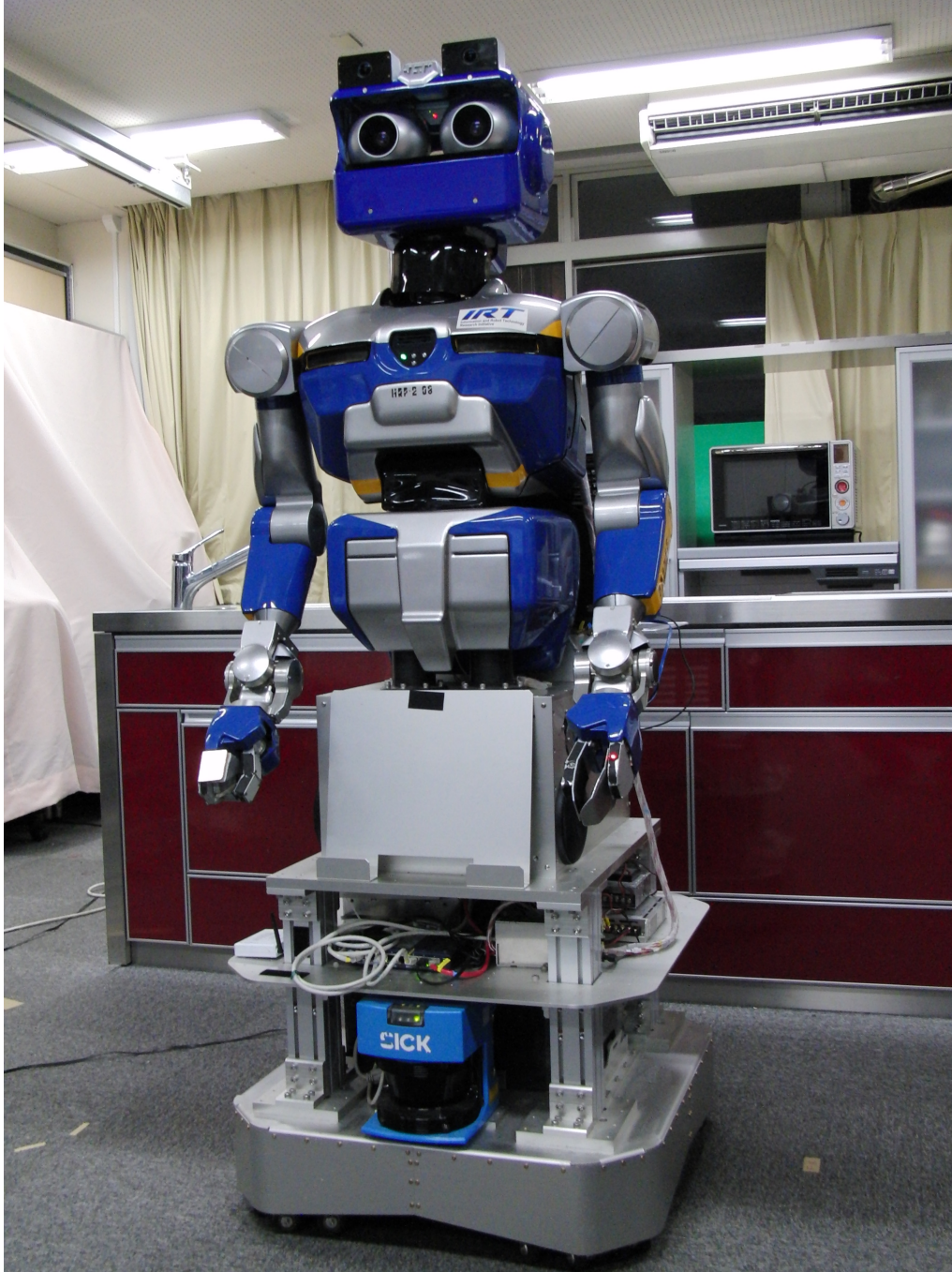


図 3.1: 広視野・高視力・等身大ヒューマノイドロボット“HRP-2VZ”（長濱・白山 2010, 著者論文 [99] にて初出）

ベル以下の記述を有するシステムを用いる．対象目標レベルの記述には，幾何学モデリング機能を有するオブジェクト指向言語である euslisp[61, 116] を用い，後述の逆運動演算により動作生成をおこなった．また，本研究6章で用いる作業目標レベルの計画器としては，PDDL (Planning Domain Definition Language) [28] を用いるものとした．各 action に相当する動作は euslisp 関数として記述し，タスクプランニング結果に従って動作を生成していく．

そして本研究では，ロボットの動作生成に，微分逆運動学計算を用いたマニピュレーション軌道の計画法 [117] を用いる．エンドエフェクタの位置および姿勢 r は，関節角度ベクトル θ を用いて，

$$r = f(\theta) \quad (3.1)$$

と書ける．しかし，この逆関数 f^{-1} は，一般に非線形な関数となる．

$$\theta = f^{-1}(r) \quad (3.2)$$

そこで，(3.1) を時刻 t について微分し，速度・角速度ベクトル \dot{r} に関する以下の式を得る．

$$\dot{r} = \frac{\partial f}{\partial \theta}(\theta) \dot{\theta} \quad (3.3)$$

$$= J(\theta) \dot{\theta} \quad (3.4)$$

ただし $J(\theta)$ は， $[r$ の次元] \times $[\theta$ の次元] 列のヤコビ行列である．

ヤコビ行列 $J(\theta)$ が正則であるときは逆行列 $J(\theta)^{-1}$ が計算でき，関節角速度 $\dot{\theta}$ は以下のように求めることができる．

$$\dot{\theta} = J(\theta)^{-1} \dot{r} \quad (3.5)$$

しかし $J(\theta)^{-1}$ は一般に正則でないため，ヤコビ行列の擬似逆行列が用いられることもある．しかしこの手法では， $J(\theta)^{-1}$ がフルランクでなくなる特異点に近づいた場合，関節角速度 $\dot{\theta}$ が大きくなり，危険な動作が計画されてしまう．そこで，Nakamuraらが提案した SR-Inverse [55] を用いる．SR-Inverse $J^*(\theta)$ は，

$$J^*(\theta) = J^T (J J^T + \epsilon E_m)^{-1} \quad (3.6)$$

のように表される。本研究では、関節角度限界を回避するとともに [29, 19] ロボットのリンク同士の衝突も回避する [65, 64] ための手法を併用して動作生成をおこなう [117] が、本研究の扱う範囲ではないため、詳細は割愛する。

HRP-2VZ に道具を用いた動作を生成させる際には、文献 [123] で述べられているような、道具の把持に関する情報 (handles) と道具を作用させる部位 (attentions) を有するような道具の幾何モデルを準備し、マニピュレータの軌道を計算させるものとした。

3.4 人間と同等の視力を有するロボットヘッドの構成法

ロボットの視力を上げる、つまり一般に利用されるカメラの角度分解能を小さくする手法としては、望遠レンズを利用する方法と、高解像度のセンサを用いる方法が存在する。このうちセンサの画素数や画素あたりのサイズには限界があることから、望遠レンズを用いたカメラを用い、視野の狭さを克服するために広角カメラを組み合わせた複合ビジョンシステム [4]、そして人間の中心窩の構造を模したレンズを使用する方法 [45] などが実証されてきている。

本章では、著者が修士論文研究および博士論文研究において利用した、等身大ヒューマノイド用・広視野高視力複合ビジョンヘッド (JSK 内での通称はズームヘッド、英名 CV-Head [51]) について、その開発史とハードウェア・ソフトウェア構成を述べるとともに、複合ビジョンヘッドシステムの長所と短所について考察する。さらに、この短所を克服するための新ヘッドについても説明する。

3.4.1 ズームヘッドの開発史

ズームヘッドは、西野環氏が修士課程在籍中に等身大ヒューマノイド用ヘッドとして開発した、輻輳機構を有し光学ズームが可能なステレオビジョンシステム [90, 91] が元となっている。博士課程在学中の袖山慶直氏のデザインにより魚眼カメラが搭載されるとともに、博士課程在学中の小島光晴氏の見守り用ロボット “Mamoru” [110] 用視覚ソフトウェアシステムを元に、修士課程在学中 (当時) の白山翔太氏と植木竜佑氏によりソフトウェアシステムが構築された [152]。並行して白山氏によって輻

轆ズーム機構を活用したアプリケーション研究が行われていたが、ステレオキャリブレーションがおこなわれていなかったため、単眼の画像処理をベースとするものであった [150, 151, 153]。さらに初号機（青色外装¹）に加えて、二号機（赤色外装²）が組み立てられた。白山氏が自身の修士論文研究として、輻轆機構とズーム機構双方を備えるステレオカメラ系のカメラキャリブレーション手法を実装したが、そのカメラキャリブレーション [149, 51, 52] および HRP-2V の頭部換装より、著者も加わった。その後、著者が修士課程および博士課程在学中、山崎公俊氏の助言を元に、同じく修士課程在学中であった西野友博氏とともにキャリブレーション手法およびハードウェア・ソフトウェアシステムの改良をおこない、著者は人や道具の追跡機能 [51, 52, 105, 107] および家事支援に必要な筒状形状認識機能 [99] を、西野氏は生活空間内の文字領域の発見および文字の読み取り機能 [142, 143] を、ズームヘッドを用いて各々実現した。さらに著者が博士研究において、重畳関係を推定し利用する視覚機能の実現研究 [53, 101, 102, 106] に活用したものである。

この章の記載だけでなく、ズームヘッドを用いた実験は多くの皆様の成果をまとめたものであることを記し、ズームヘッドの記述へと入る。

3.4.2 ズームヘッドのハードウェア構成

ズームヘッドのカメラ位置および自由度配置を、図 3.2 に示す。ズームヘッドは、光学ズーム可能なカメラを二基、魚眼レンズを搭載したカメラを二基、搭載している。光学ズームレンズを搭載したカメラとしては Sony 社の “FCB-IX11A” であり、ズーム倍率は 1.0 から約 9.5 まで変化させることが可能である。0.7× ワイドコンバージョンレンズ “VCL-HG0730X” を装着することで、取り得る水平画角は 65.7[°] から 6.9[°] までとした。VGA (640[pixel] × 480[pixel]) サイズの画像を用いた場合、TELE 端での水平角度分解能は

$$65.7[^\circ] \div 640[\text{pixel}] = 1.03 \times 10^{-2}[^\circ/\text{pixel}] \quad (3.7)$$

であるから、最小分解能の逆数である視力は約 1.6 相当となる。二基のカメラはマクソン社の AC 40[W] モータと、ギアレシオ 300:1 のハーモニックギアを用い、パン軸を

¹白山氏による愛称は「青ヘル」

²白山氏による愛称は「赤ヘル」

独立に駆動することが可能である．パン軸の関節角度センサとしては Microtech Laboratory 社のロータリエンコーダ “MEH-30-4500 PST16 (72000)” を採用した．300:1 のハーモニックギアを解すると，エンコーダの最小角度分解能は $0.00125[^\circ/\text{count}]$ となり，これはカメラの角度分解能より十分に小さくなる．また，全方位カメラとして，オプト社の “NM33” を二基用いた．NM33 は魚眼レンズを搭載し，最大画角約 $180[^\circ]$ の全方位画像を撮影できるだけでなく，任意の画角での展開画像を同時に取得することが可能である．ズームヘッドを HRP-2V へ搭載した際のハードウェア・ソフトウェア構成を図 3.3 に示す．HRP-2V に搭載されることで，ロボットの身長は約 $166[\text{cm}]$ と人間並みとなり，その視点は人間の視点と同等なものとなることから，人間と同様の物体の見え方，特に人間と類似した視覚重畳関係を観測可能な視覚となることが期待できる．

3.4.3 ズームヘッドのソフトウェア構成

輻輳ズームステレオシステムでは，ズームカメラのズーム倍率変化によってカメラの内部パラメタが変化する．また，輻輳角度の変化により外部パラメタが変化する．そこで本システムでは，視覚処理ループ毎に輻輳角度とズーム倍率を取得し，内部パラメタおよび外部パラメタの再計算をおこなう構成とした [149]．

まず，内部パラメタの決定法について述べる．本システムでは，カメラモデルとして一般的なピンホールカメラモデル（文献 [12] 等に詳しい）を仮定する．このモデルでは，カメラ座標系での三次元点 (X, Y, Z) が，内部パラメタ行列 A を有するカメラにより，画像平面上の二次元点 (u, v) へ投影される場合，それらの関係は同次座標系を用いて以下のように表される．

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.8)$$

ただし内部パラメタ行列 A は，焦点距離 $f_x[\text{pixel}]$ ， $f_y[\text{pixel}]$ および主点 $(c_x[\text{pixel}],$

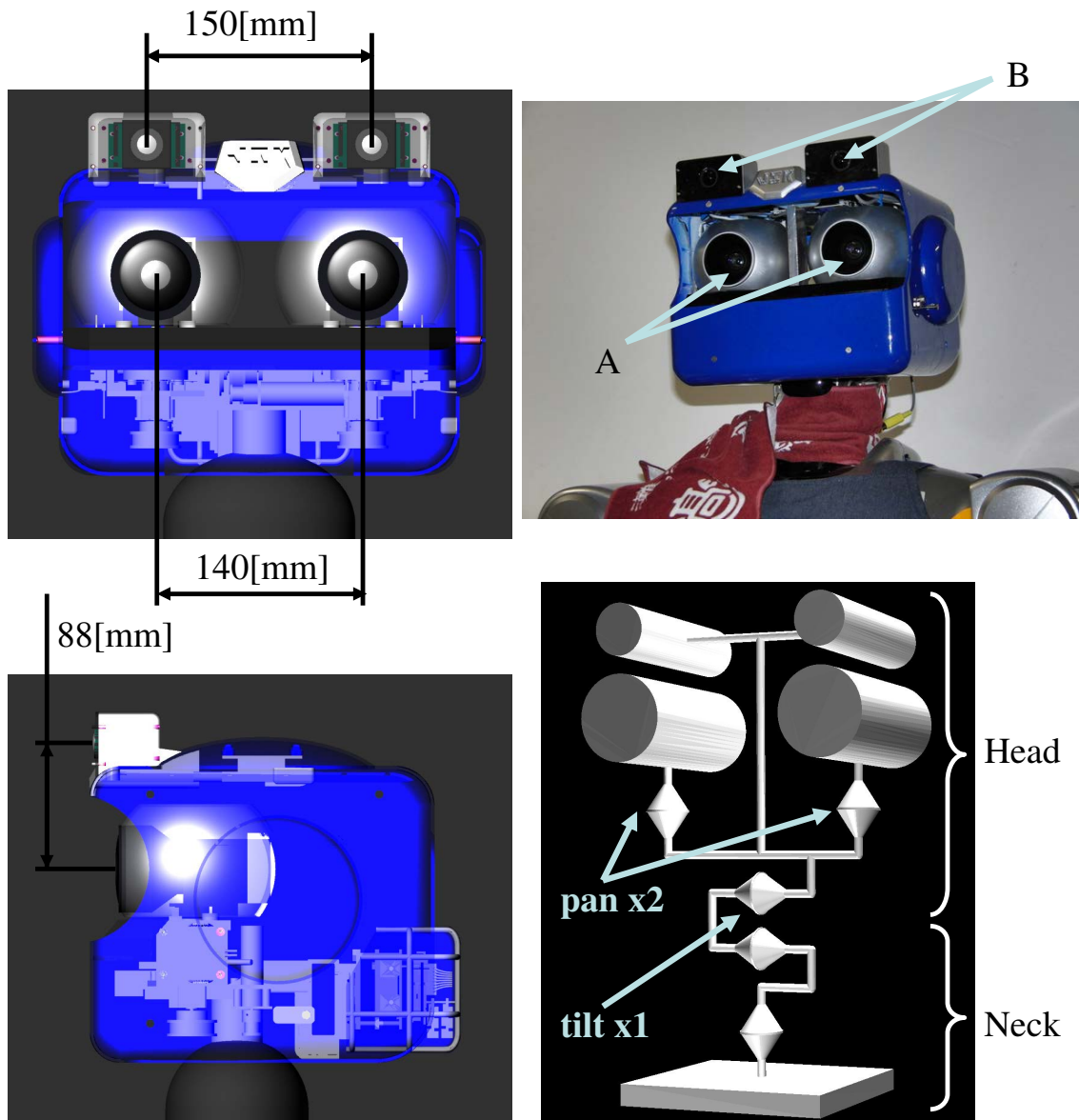


図 3.2: ズームヘッドの寸法概要・自由度配置 (著者論文 [52] より)

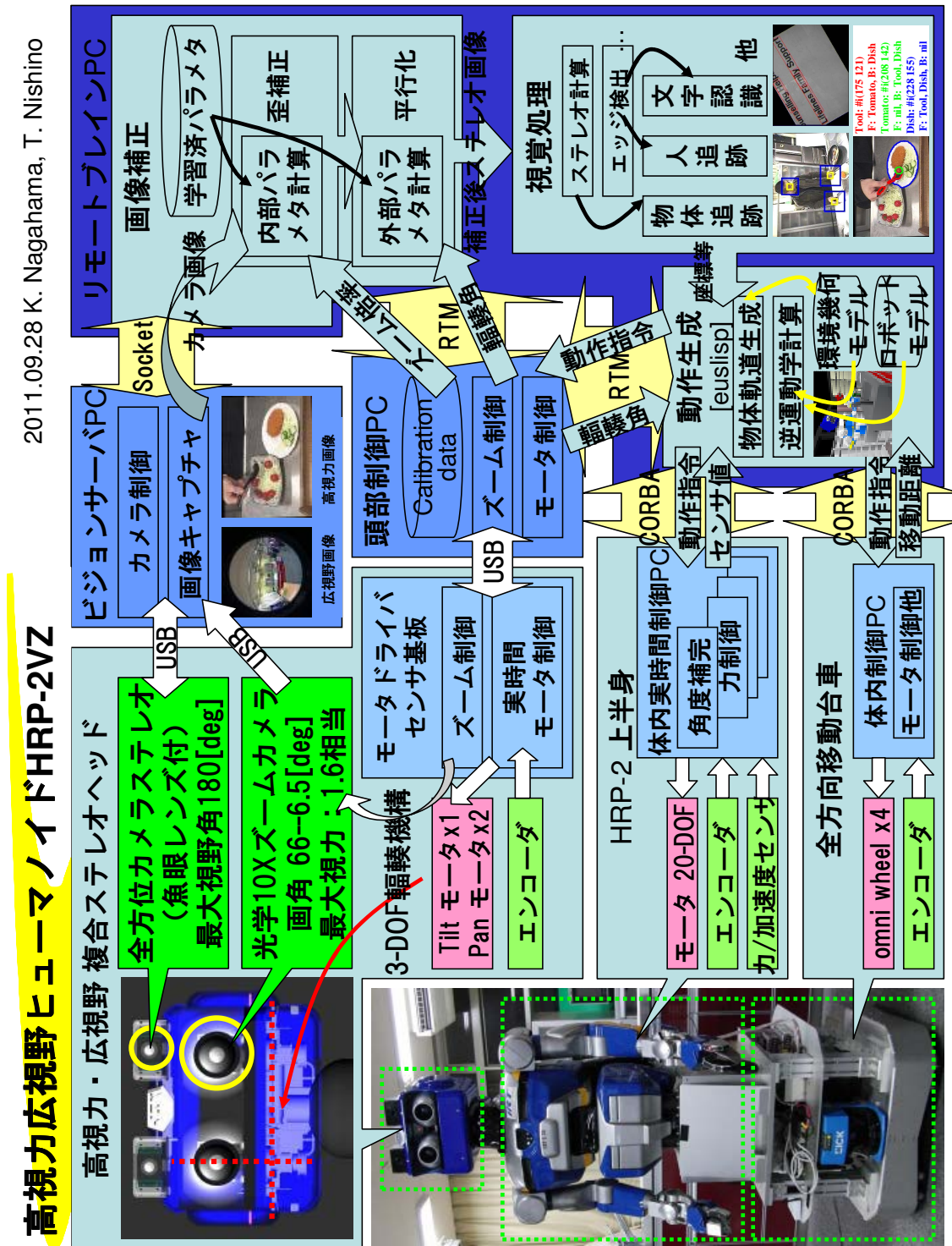


図 3.3: HRP-2VZ のハードウェア・ソフトウェア構成 (長濱・西野 2011, 西野修士論文 [142] にて初出)

c_y [pixel]) を用いて、以下のように表される。

$$A = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3.9)$$

しかし実際のレンズには様々な収差が存在するため、位置補正式を用いて画像を補正するのが一般的である。レンズ形状が理想的な双曲線からずれていることで発生する半径方向歪み、そしてレンズと画像素子の理想的な位置関係からのずれで発生する円周方向歪みは、Brown らが提案した次式 [13, 14] により補正できることが知られている。 (x', y') を補正前画像内位置、 (x, y) を補正後画像内位置として、

$$x' = x + \bar{x}(K_1 r^2 + K_2 r^4 + K_3 r^6 + \dots) + [P_1(r^2 + 2\bar{x}^2) + 2P_2 \bar{x} \bar{y}][1 + P_3 r^2 + \dots] \quad (3.10)$$

$$y' = y + \bar{y}(K_1 r^2 + K_2 r^4 + K_3 r^6 + \dots) + [2P_1 \bar{x} \bar{y} + P_2(r^2 + 2\bar{y}^2)][1 + P_3 r^2 + \dots] \quad (3.11)$$

ただし、 K_i と P_i は定数、 (x_p, y_p) は主点を表し、

$$\bar{x} = x - x_p \quad (3.12)$$

$$\bar{y} = y - y_p \quad (3.13)$$

$$r = \sqrt{(x - x_p)^2 + (y - y_p)^2} \quad (3.14)$$

とする。本システムではこのうち、半径方向歪みを補正する一次項のみを利用するものとし、以下の式により歪補正をおこなった。ただし (x, y) を主点を原点とした補正前画像内位置、 (x_c, y_c) を主点を原点とした補正後画像内位置とし、 k_1, k_2 は定数である。

$$x = x_c \{1 + k_1(x^2 + y^2)\} \quad (3.15)$$

$$y = y_c \{1 + k_2(x^2 + y^2)\} \quad (3.16)$$

これは円周方向歪みが十分に少ないカメラ系であったからである。

次に、本システムではズーム倍率変化に応じて変化する内部パラメタを、焦点距離 f_x, f_y 、歪みパラメタ（半径方向歪みパラメタ k_1, k_2 ）および主点 (c_x, c_y) と仮

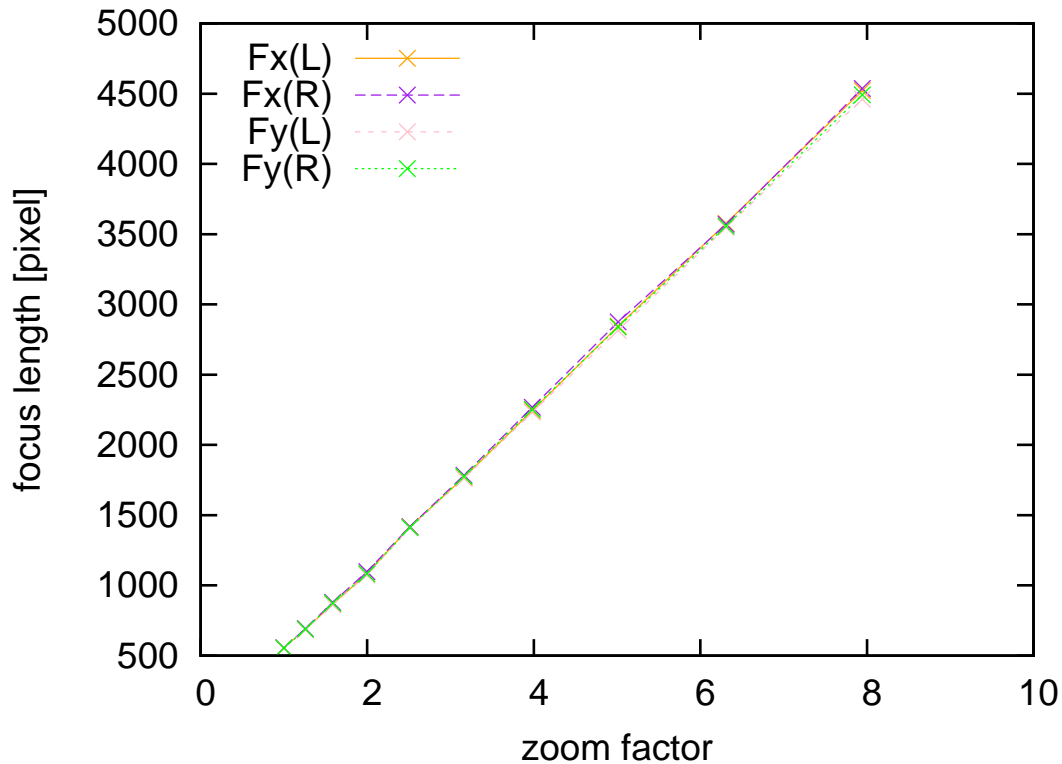


図 3.4: ズーム倍率と推定された焦点距離の関係 (長濱・白山 2010, 著者論文 [51] より, 改変前の初出は [148])

定した．主点はズーム中心を使用するものとし，固定したチェスボード中の四つの特徴点について，ズーム倍率を変化させた際の画像内軌跡を求め，それらの軌跡の交点を採用した．

f_x , f_y 及び k_1 は，予めズーム倍率 $10^{i/10}$ ($i = 0, 1, 2, \dots, 9$) [倍] について，Zhang のカメラキャリブレーション [83] により推定しておき，再計算時には得られたズーム倍率から，線形近似により求めるものとした．各々のパラメタがズーム倍率に対し線形性を示すことは，基礎実験により予め確認した (図 3.4) ．

輻輳軸を有するカメラの外部パラメタには，左右それぞれのカメラと輻輳回転軸の幾何関係，及び左右の回転軸の幾何関係が存在する．Welke らは，輻輳軸を回転させながら，カメラで固定したチェスボードを観測し，回転軸とカメラの幾何関係

を推測した [73] . 本研究でもこの手法を用い , 左右カメラの回転軸位置関係を保持しておき , 再計算時には , 得られた左右輻輳軸角度に基づき外部パラメタを計算するものとした [149] .

3.4.4 広視野視覚と高視力視覚を併用した注視制御

提案したズームヘッドでは , 光学ズームカメラが狭角画像を取得している際にも , 全方位カメラを併用することで人の体全体の追跡および , それを用いた注視制御をおこなうような協調制御が可能である . 本節ではこれを検証するため , 二基の全方位カメラを画角約 90° とした広角カメラとして用い , 著者らが開発した連続視差フィルタ [52] を用いて人の頭部および両手先を追跡しながら , 光学ズームカメラのズーム倍率と頭部姿勢を変更していく実験をおこなった .

連続視差フィルタは , ゼロ視差及びその周辺の特徴抽出を行う拡張ゼロ視差フィルタ [72] を , 注視点並列制御を可能とするためにゼロ視差以外へ拡張したものであり , 時刻 $T[\text{frame}]$ での三次元点存在位置 $p(T)$ と時刻 $T+1[\text{frame}]$ のステレオ画像を入力とし , 三次元形状が連続性を示す画像内候補点 p_I と , そのカメラ座標系での三次元位置 $p(T+1, i)$ を出力する . 手法の詳細は [52] を参照されたい .

本実験では , 広視野視覚の左目画像内で , 両手の追跡結果がどこに存在するかを評価し , 高視力視覚の制御に利用するものとした . 図 3.5 に詳細を示す . 例えば , 図 3.5 の左図において , “Up” と書かれた領域に左右いずれかの手の追跡結果が含まれた場合 , 首部のピッチ軸を 0.5° 上へ向ける . “Down & Right” と書かれている領域では , 首部下及び右へそれぞれ 0.5° だけ動かす . ただし , 左右の手の追跡結果が正反対の動作を指示した場合は , 動かさないものとした . 図 3.5 の右図において , 左右両方の手の追跡結果が中心部に含まれた場合には , ズーム倍率を $0.04[\text{倍}]$ 大きくする . そのほかの場合は $0.04[\text{倍}]$ 小さくするものとした . ただし , ズーム倍率は $1.0[\text{倍}]$ から $2.0[\text{倍}]$ までとした .

広視野視覚により人の上半身を追跡した結果および高視野視覚で撮影された画像を図 3.6 に示す . 広視野視覚では人の頭と両手の追跡に成功し , 一フレームあたり三点を追跡するために要した時間は $30[\text{milisecond}]$ 程度であった . 次に , 実験中のズーム倍率 , 頭部ヨー軸角度およびピッチ軸角度の推移をそれぞれ図 3.7 , 図 3.8 ,

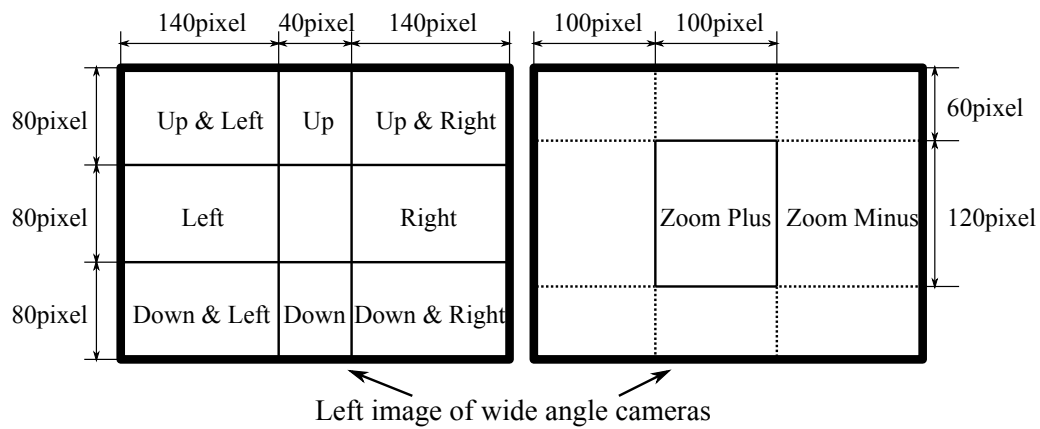


図 3.5: 広角カメラでの人の上半身追跡結果を用いた, 高視力視覚の誘導方法 (著者論文 [52] より)

図 3.9 に示す. 高視力視覚はズーム倍率が大きくなると同時に, 注目点が下へ下がっている. これにより高視力視覚と広視野視覚を統合した注視制御が可能であり, 人の両手へ分解能の高い注意を向けることが可能であることが示された.

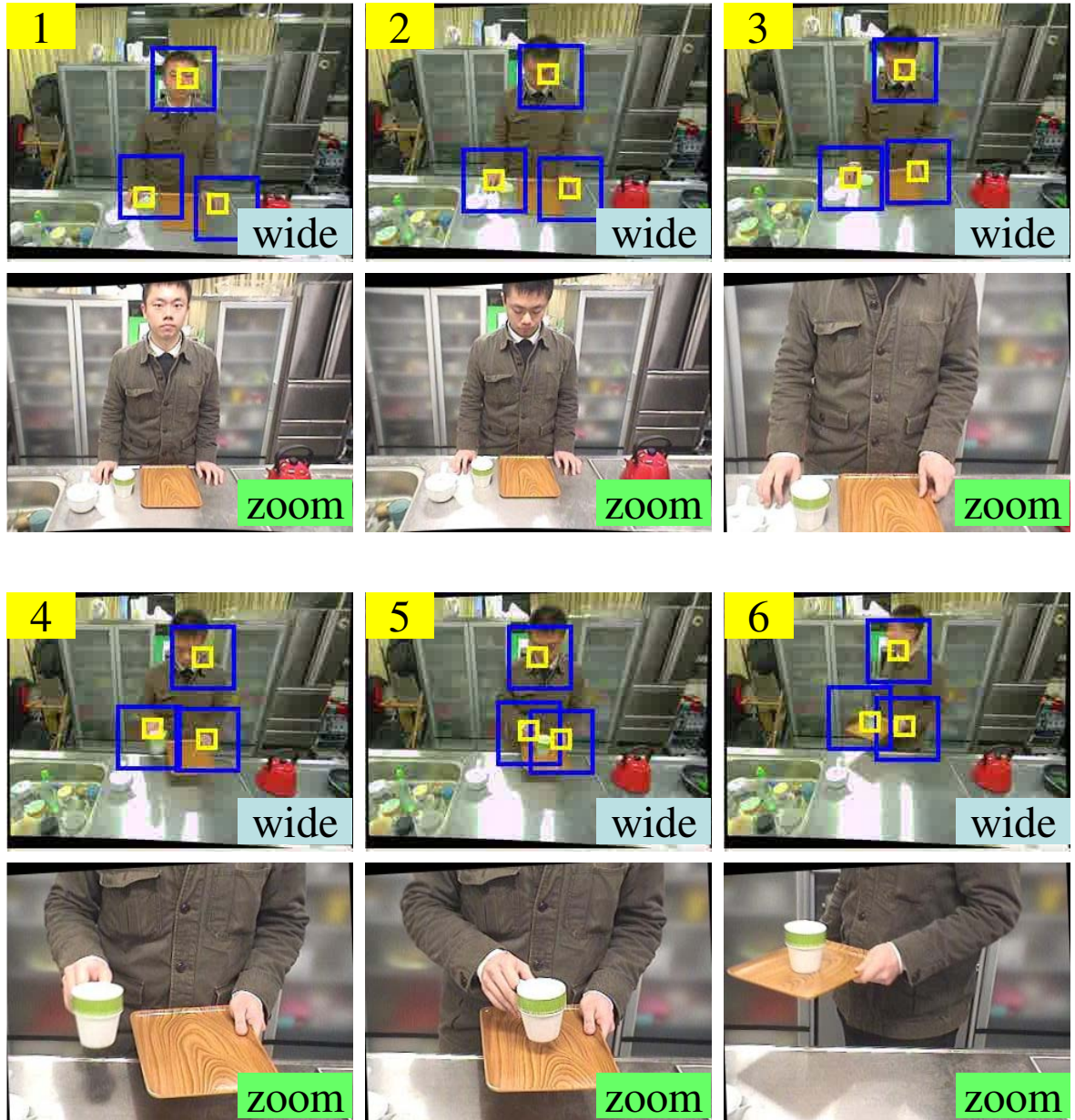


図 3.6: 高視力視覚と広視野視覚の協調実験結果 (著者論文 [52] より)

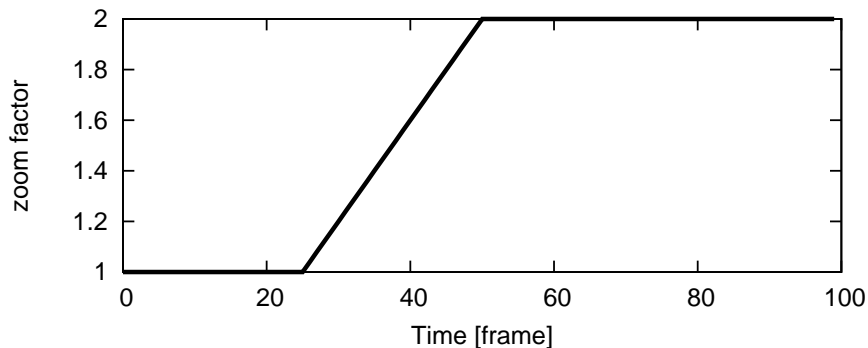


図 3.7: 高視力視覚と広視野視覚の協調実験における，ズーム倍率の推移（著者論文 [52] より）

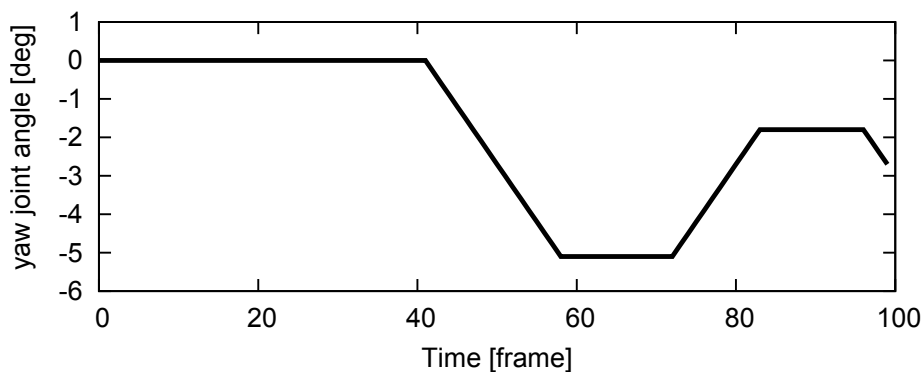


図 3.8: 高視力視覚と広視野視覚の協調実験における，首部ヨー軸角度の推移．負の値は，左を向いていることを示す（著者論文 [52] より）

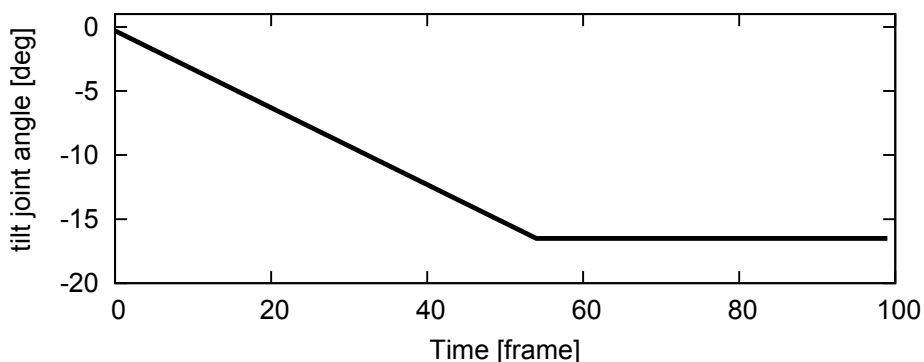


図 3.9: 高視力視覚と広視野視覚の協調実験における，首部ピッチ軸角度の推移．負の値は，下を向いていることを示す（著者論文 [52] より）

3.4.5 輻輳ズームステレオ機構に関する考察

ズームヘッドを制作するにあたって、輻輳ズームステレオ機構の利点としては、以下の点が考えられた。

高解像度かつノイズの少ない画像の取得

光学ズームレンズを搭載したカメラに起因する利点である。撮像素子数に限界があったとしても、光学ズームを用いて望遠画像を取得すれば、注視点周辺の高解像度画像を劣化なく取得できる。また望遠画像はレンズの中心部を利用することになるため、歪みに起因する収差の影響が小さくなることが期待できる。

転送画像サイズの縮小

光学ズームレンズを搭載したカメラに起因する利点である。画像処理に利用したい注視点周辺にズームし、局所画像のみを画像処理用計算機へ転送し利用するようにすれば、広角の高解像度画像よりも高速な転送を期待できる。

高速な視線方向移動の実現

ロボットの頭部へ、カメラ以外にも様々なセンサ類を搭載する場合、その重量は相当なものとなる。視線方向を変化させる際に、重い頭部全体を動かすのではなく、カメラ単体のみを動作させることができれば、高速な視線方向移動が可能だと考えられる。

教示者への注視点のフィードバック

人間がロボットの視線方向を知ることができれば、ロボットへ何かを教えるシーン等で、正しいインタラクションが取れているか否かを知るための手がかりとなり得る。輻輳機構を有するロボットが、注視対象を左右それぞれのカメラの中央で捉える場合、首のみに自由度を有するロボットとは異なり、注視点の距離に関する情報を人間へフィードバックできる可能性がある。輻輳機構を有するアンドロイドの視線方向を人間が推定した際に、アンドロイドが輻輳軸を動かした方が真値に近い推定が可能だとの知見もある [88]。

しかし運用上、輻輳ズームステレオ機構には、以下の短所の存在が認められた。

カメラ外部パラメタ再計算結果の誤差

輻輳機構を有するカメラシステムに起因する短所である．外部パラメタの再計算においては，1) 推定された二つの輻輳軸同士の位置・角度関係，2) 各々の輻輳軸からカメラまでの距離・角度，そして3) 各々の輻輳軸回りの回転角度が必要となる．主にカメラが搭載されたリンクとエンコーダの機械的なずれに起因して，3) に誤差が生じることが多かった．また，その誤差が存在する状態で1) と2) を計算しているため，二つのカメラ同士の位置関係には誤差が生じることとなり，外部パラメタの正しさは保証されなかった．

また，キャリブレーションの煩雑さも大きなデメリットの一つであった．

配線の可動部で生じるノイズ

輻輳機構を有するカメラシステムに起因する短所である．カメラと計算機を有線ケーブルで結合している場合，輻輳可動軸に近い部分で，かつ柔軟に引っかかりなく配線をおこなわなければ，ケーブルへの張力変化によってノイズが乗る，ケーブルや接続部が破損する等の問題が生じた．

頭部重量に起因する構造上の問題

カメラだけでなく，輻輳機構のためのモータや減速機等の重量が頭部に集中するため，首部構造の補強および高出力モータへの換装が必要であった．構造の補強後も，その限界から剛性不足に起因する頭部姿勢の誤差が存在した．

その他，輻輳軸の最大関節角速度は $200[^\circ/\text{second}]$ 程度であったが，このような速度で駆動しながら画像を取得した場合，1) 画像取得タイミングと関節角度取得タイミングのずれにより，外部パラメタ推定に誤差が生じる；2) ズームヘッドへ使用したカメラの特性上，画像に大きなブレが生じるという問題もあった．利点で述べた高速な視線方向移動の実現のためには，より高速な画像取得と関節角度センサとの時刻合わせが重要な点となることが分かった．

3.4.6 ズームヘッド改良版 “CV-Head Mk.II”

前節で考察した課題を解決するために，新たな等身大ヒューマノイド用ヘッド “CV-Head Mk.II” (図 3.10) を開発し，併用するものとした．CV-Head Mk.II では，輻

鏡機構に起因するカメラ外部パラメタ再計算の誤差に関する問題を解決するために、高解像度カメラ二基を固定ステレオとして用いるとともに、距離カメラを導入して距離計測の精度を向上している。このヘッドはHRP-2Vへ搭載し、5章と6章での観察実験に用いた。

3.4.7 CV-Head Mk.IIのハードウェア・ソフトウェア構成

左右の高解像度ステレオカメラとして使用しているのは、Baumer社製の産業用 Gigabit Ethernet カメラ “TXG20c-P” [8] である。以後、GigE カメラと呼ぶ。センサーサイズは $1/1.8''$ ($6.9[\text{mm}] \times 5.2[\text{mm}]$)、解像度 $1624[\text{pixel}] \times 1232[\text{pixel}]$ 、最大フレームレート $16[\text{fps}]$ であり、Power over Ethernet (PoE) 技術により LAN ケーブルを用いた給電が可能である。GigE カメラには $F1.2/6\text{mm}$ の単焦点レンズを組み合わせ、水平画角は約 60° となっている。センサの最小分解能は

$$60^\circ \div 1634[\text{pixel}] = 3.7 \times 10^{-2} [^\circ/\text{pixel}] \quad (3.17)$$

であるから、最小分解能の逆数である視力は約 0.45 相当となる。この視力はズームヘッドの光学ズームカメラの最大視力 1.6 相当より約 3 倍劣るが、単焦点レンズであることに起因する画質の良さ、そして外部パラメタの再計算誤差が無いことを考慮すると、光学ズームカメラに匹敵する性能を有すると考えた。

距離カメラとして用いたのは ASUS 社製の RGBD カメラ、“Xtion PRO LIVE” [6] である。システムではフレームレート約 $30[\text{fps}]$ で、RGB 画像、反射赤外光を用いた距離画像ともに $640[\text{pixel}] \times 480[\text{pixel}]$ サイズの画像を、ほぼ時間のずれもなく取得できている。

本システムをズームヘッドと比較した場合の利点としては、以下のものが挙げられる。

構造の簡潔さに起因する利点

ズームヘッドより自由度が少ないため、必要な関節キャリブレーションが少なくなるだけでなく、配線の取り回しが簡潔となり、ノイズや故障の原因を削減することができる。

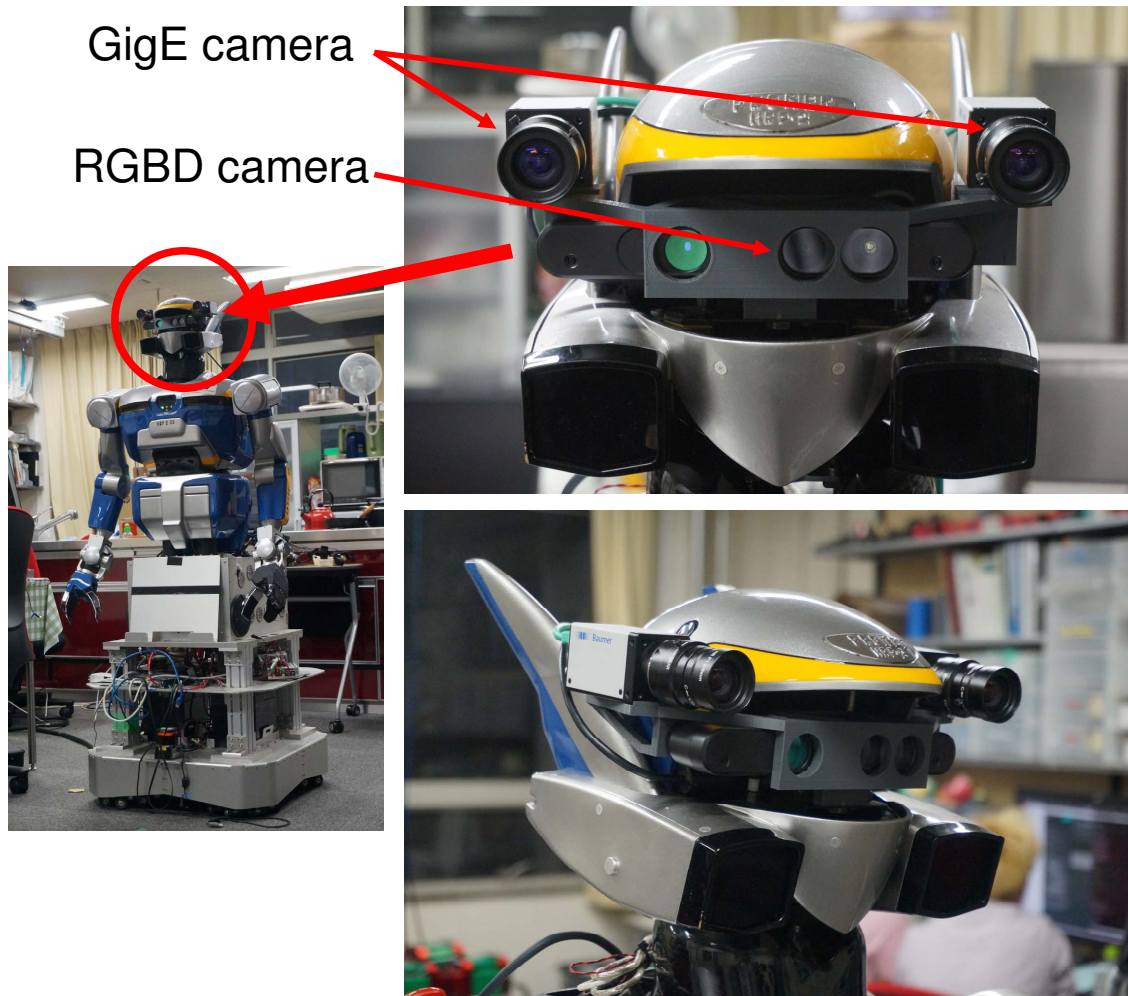


図 3.10: ズームヘッド改良版 “CV-Head Mk.II”

カメラ内部パラメタ・外部パラメタの正確性

本システムにおいては、ロボットの動作に付随して光学系が変化する状況は、視点移動以外には存在しない。そこで、各カメラの内部パラメタ推定結果、およびカメラ間の位置関係（外部パラメタ）の推定は予め一度おこなうのみとし、固定値を用いることができる。ズームヘッドで生じていた、パラメタ再推定に起因する問題が発生しないため、常に高精度で位置合わせに成功していることが期待できる。

距離情報の正確性

ステレオによる三次元再構成と比較して、反射赤外光を利用した距離カメラでは、テクスチャの少ない対象に対しても、安定した距離情報を得ることができる。

RGBD カメラの RGB 画像と距離画像の間の位置合わせは成功しているものとし、RGBD カメラの RGB 画像と GigE カメラの画像を用い、外部パラメタ推定をおこなって両カメラの位置合わせをおこなった。外部パラメタ推定後、RGBD カメラ画像を GigE カメラ（左）視点に合わせて重畳させた結果を図 3.11 に示す。二つの画像が一致していることから、外部パラメタ推定の正当性が確認できた。

3.5 HRP-2VZ による道具利用動作の考察

本節では、ロボット自身の道具利用に必要な視覚・動作機能の考察のため、HRP-2VZ に複数の道具を使わせる基礎実験をおこなった。この基礎実験では HRP-2VZ に、目玉焼きを焼いて皿によそいつける、一連の調理支援タスクをおこなわせた。実験の様子を、図 3.12 から図 3.17 に示す。各々の図において、ロボットの動きを外部カメラで観測した結果を図中 A に、euslisp シミュレーション環境でのロボットモデル、つまりロボット姿勢の指令値を図中 B に示している。A と B の後ろの番号は時刻順となっており、同じ番号が同じ時刻に対応している。図中 A のラベルの右側の数値は、動作開始からの時刻（分・秒）を示す。

HRP-2VZ の作業手順は以下の通りである：まず、生卵が入った容器をつかんで持ち上げ（～ A05）、生卵をフライパンにあけて（A7～A13）、容器を元に戻す（A14～

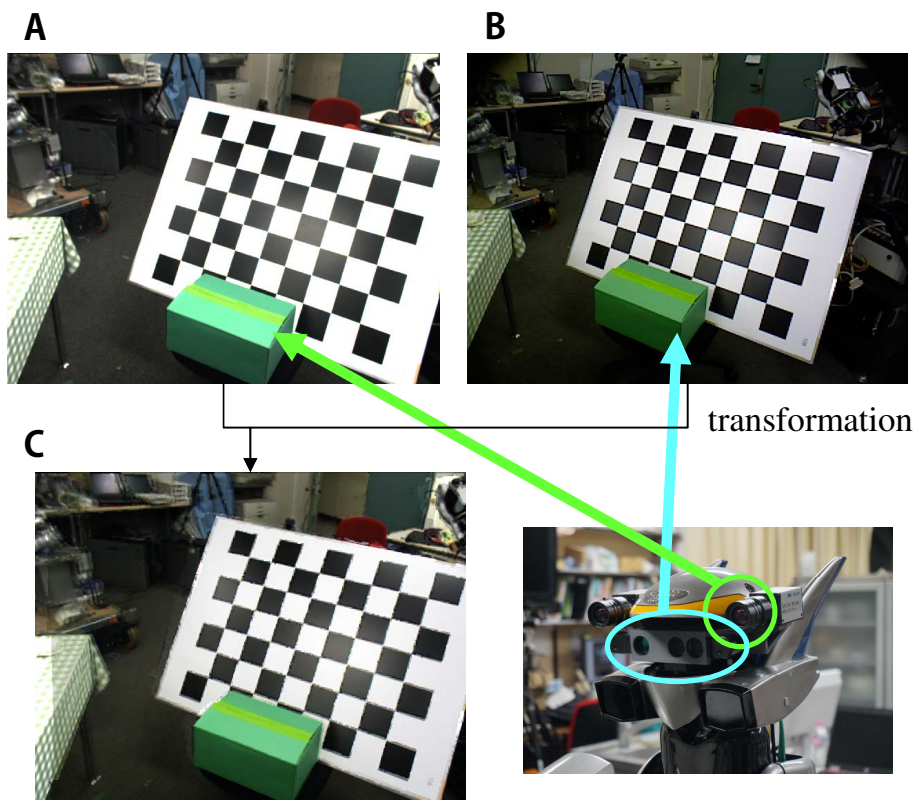


図 3.11: 高解像度 GigE カメラと RGBD カメラの外部パラメタ推定に基づく位置合わせ結果 . RGBD カメラ画像を GigE カメラ視点に合わせて重畳表示している

A19)．次に蓋を持ち上げて (A20～A23)，フライパンにセットする (A24～A27)．目玉焼きが焼けるまで3分待ち (A28～A29)，蓋を開けて (A30～A33)，蓋を元の場所に戻す (A34～A37)．さらにフライパンがずれないように柄をつかんで (A38～A39)，フライ返しに目玉焼きを載せ (A40～A42)，目玉焼きを持ち上げ (A43)，皿へよそいつける (A44～A48)．これらの操作には，本章のはじめで述べたマニピュレーション能力が必要となり，HRP-2VZ が等身大かつ双腕であることで実現可能となっている．

基礎実験では目玉焼きの調理からよそいつけまで成功した試行も存在したことから，液体や柔軟物が操作対象となるタスクも，道具を用いて遂行可能であることが示された．一方で，いずれかの操作で失敗する試行も存在し，問題点としては以下のような点が挙げられた．

ピックアンドプレイス

基礎実験中では，容器のピックアップ (A02-A05) と容器のプレイス (A15-A18) をはじめ，前述の対象操作全てに必要な技術である．ピックアンドプレイスは人間の日常生活でも最頻出の操作であることが知られている [111] が，ロボットが道具を用いてタスクを遂行する際にも同様に重要であることがわかる．お盆のような物を運ぶ道具では，道具利用の前に対象物を道具の上へ置く必要があるのはもちろん，道具自体も持ち上げて運び，再びどこかへ置くこととなる．基礎実験中のコップのような道具についても，卵を注ぎ入れる前後で同様に発生する操作である．

このような作業においては，操作対象物を確実につかみ落とさぬこと，そして操作対象物やロボット自身をも破損しないように，周囲の環境との干渉を防ぐ必要がある．実際に基礎実験中では，生卵をフライパンにあける (A7～A13) 際に，容器をつかむ部分が期待通りで無かった場合，フライパンの端に容器が当たり，あける動作に失敗することがあった．

そこで，このような操作をおこなう前には，ロボットに対しての操作対象物の形状および位置・姿勢を特定する必要があるが，把持前の操作対象物が遮蔽下にある場合は多くないであろうことから，一般的な物体探索手法が適用可能である．操作対象物を確実につかむためには，実機動作前のシミュレーション

による把持計画 (OpenRAVE[22] ほか), エンドエフェクタのトルク制御などが一般におこなわれている。環境や操作対象物との不必要な干渉やロボットの自己干渉防止にも, 実機動作前のシミュレーションによるマニピュレーション計画 [22], 簡単には Approach point や Departure point を用いた逆運動学計算手法 ([86] に詳しい), そしてロボット自身にハードウェア・ソフトウェア的にコンプライアンスを与えて故障を防ぐ方法が考えられる。いずれの手法も, 道具のピックアンドブレースへも同様に利用できる。

このように, 操作対象物と環境の認識を十分正確におこなえば, ピックアンドブレースは道具を用いた間接的な対象操作よりも成功率の高い実現が可能だと考えられ, 本基礎実験および本論文6章の検証実験においては, Approach point と Departure point を用いた逆運動学計算に基づくマニピュレーション手法を用いている。

道具を用いた間接的な対象操作

ロボットが物体を操作している点で, 前項“ピックアンドブレース”と同様の留意事項は重要となる上, 直接の操作ではないため, 対象の操作に失敗が容易に発生する。前述の目玉焼き作業においては, 生卵をフライパンにあける (A7~A13), フライ返しに目玉焼きを載せる (A40~A42) がこのような操作に相当する。特に後者について, フライ返しを目玉焼きに差し込む動作は静的でない動作となり, 双方の間に発生する摩擦力が試行によって変化するため, 失敗が多く発生した。

道具を用いることで多様な操作が目的となるため, 操作に必要な道具と操作対象の状態を推定しつつ, 道具が操作対象に与えている影響を正しく推定しながら操作する必要がある。さらに, 目的の操作が達成されているかを確認し, 失敗している場合には失敗復帰をおこなう必要がある。

上下・包含関係の変化

操作の前後において, 道具と操作対象, あるいは複数の操作対象間の上下・包含関係が変化する状況は多く存在する。本基礎実験中では, 生卵をフライパンにあける (A7~A13), フライパンに蓋をセットする (A24~A27), フライ返しに目玉焼きを載せ (A40~A42), そして皿へよそいつける (A44~A48)

などが相当する。基礎実験ではフライパンに蓋をセットする際に蓋の手放し方やフライパンの位置変化によって蓋がずれる，皿へのよそいつけにおいて，フライ返しと目玉焼きに発生する摩擦力の変化により，皿からはみ出してよそいつけがなされる等の失敗が発生した。

このような場合には物体同士に遮蔽関係が頻繁に発生することから，前項で指摘したように，道具が操作対象に与える影響の推定において対象の位置・姿勢の特定が難しくなる。そこで，道具や対象に遮蔽が発生しても，それらの状態を推定できるような枠組みが必要である。

各々の操作において失敗が発生したことは，前章で述べた対象同士の接触関係を変化させるような操作において失敗が起こり得ることを立証するものである。したがって，このような操作には失敗復帰あるいは動作学習が必要であると言える。そのような動作中，最も成功率が低かったのは，フライ返しに目玉焼きを載せ，持ち上げる操作であり，その成功率は29[%]であった。これは，道具を用いて操作対象を操作しながら，それらの上下・包含関係を変化させるものとなり，前述の問題点を多く含むからだと考えられる。

本論文で提案し評価する手法は，このような上下・包含関係の変化に加え，包丁が操作対象に切り込むような突入関係が変化するような状態の観測に用いるものであり，詳細は次章以降で述べる。

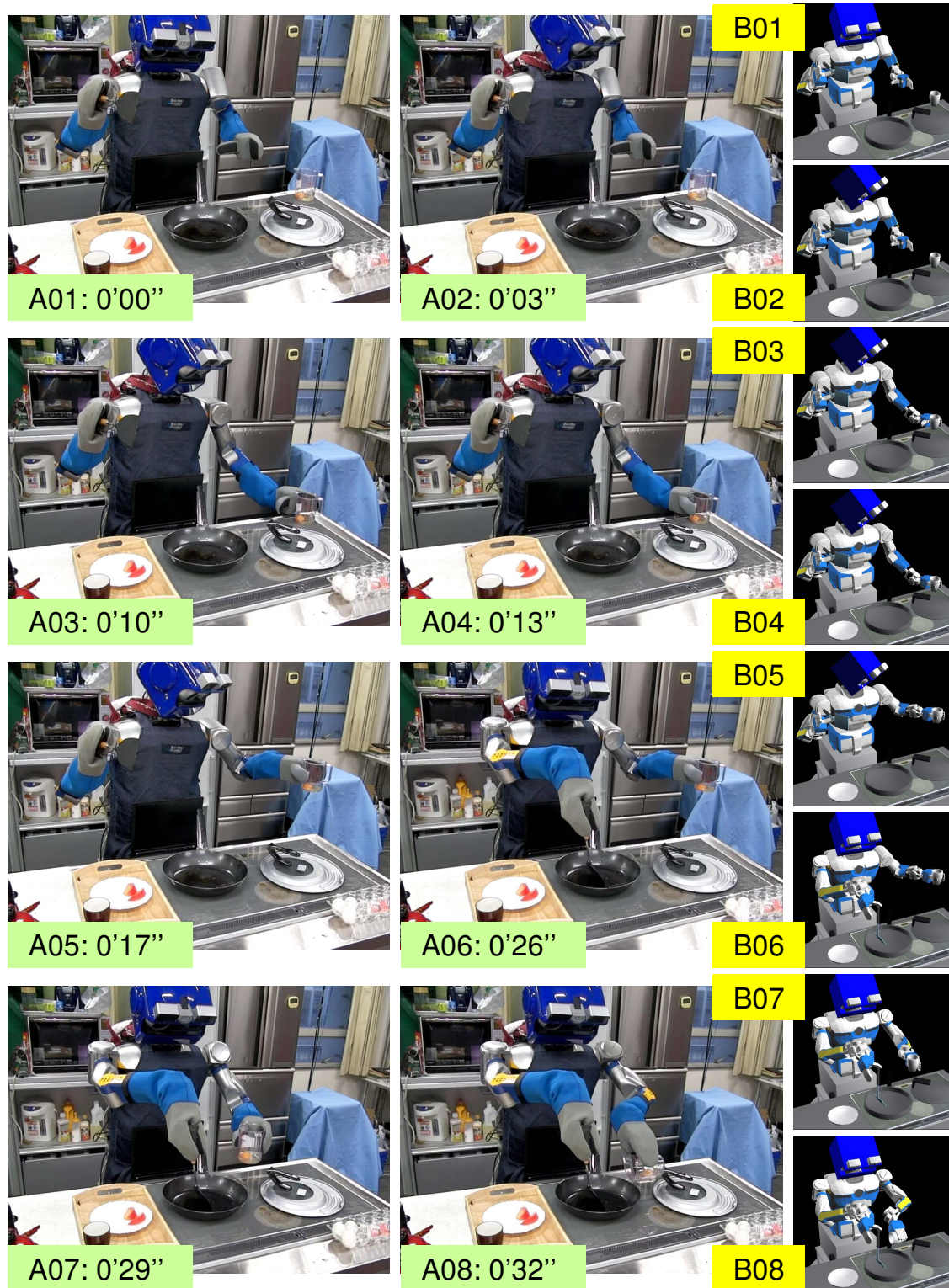


図 3.12: HRP-2VZ による目玉焼き調理実験 (1/6)

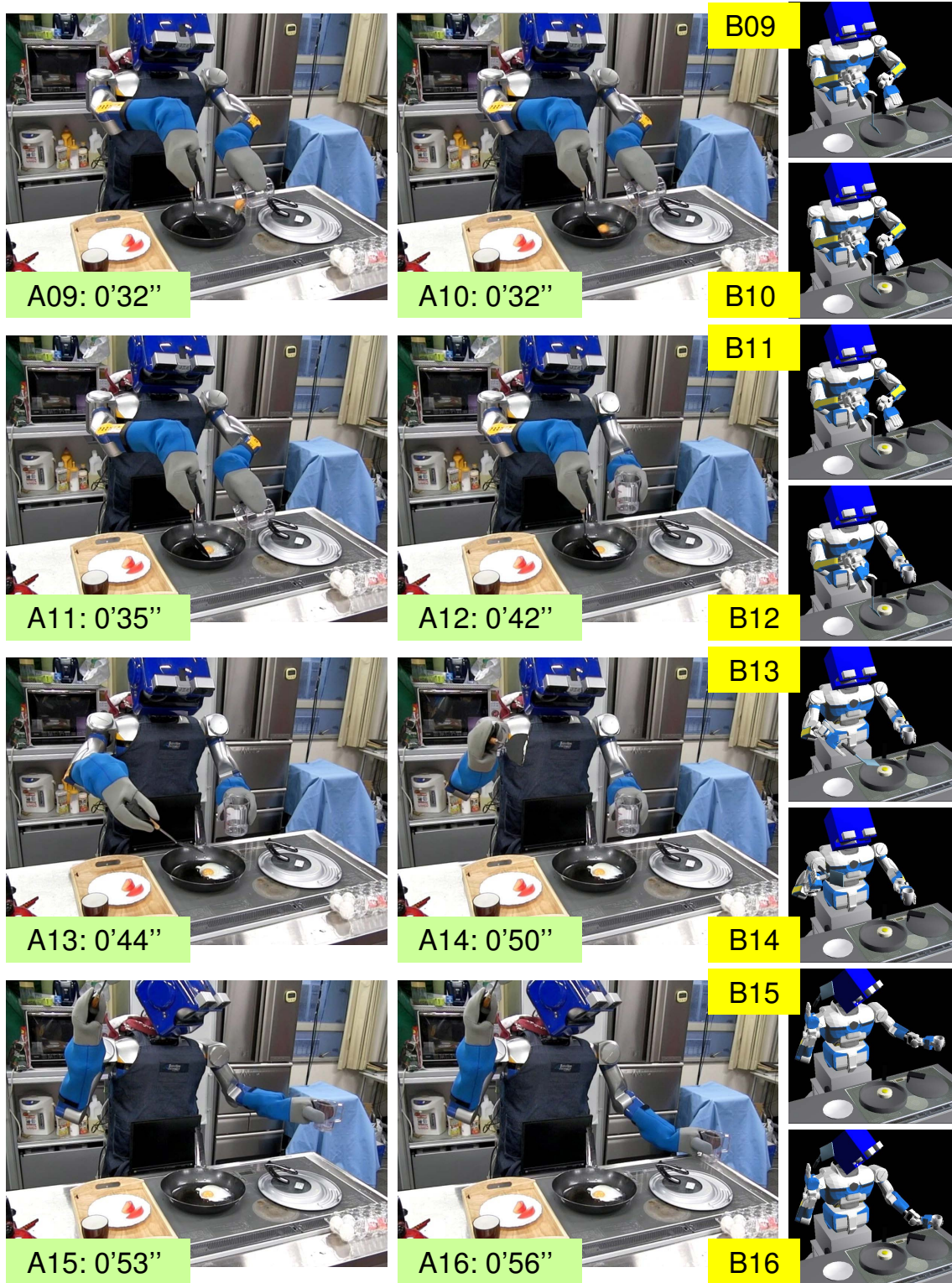


図 3.13: HRP-2VZ による目玉焼き調理実験 (2/6)

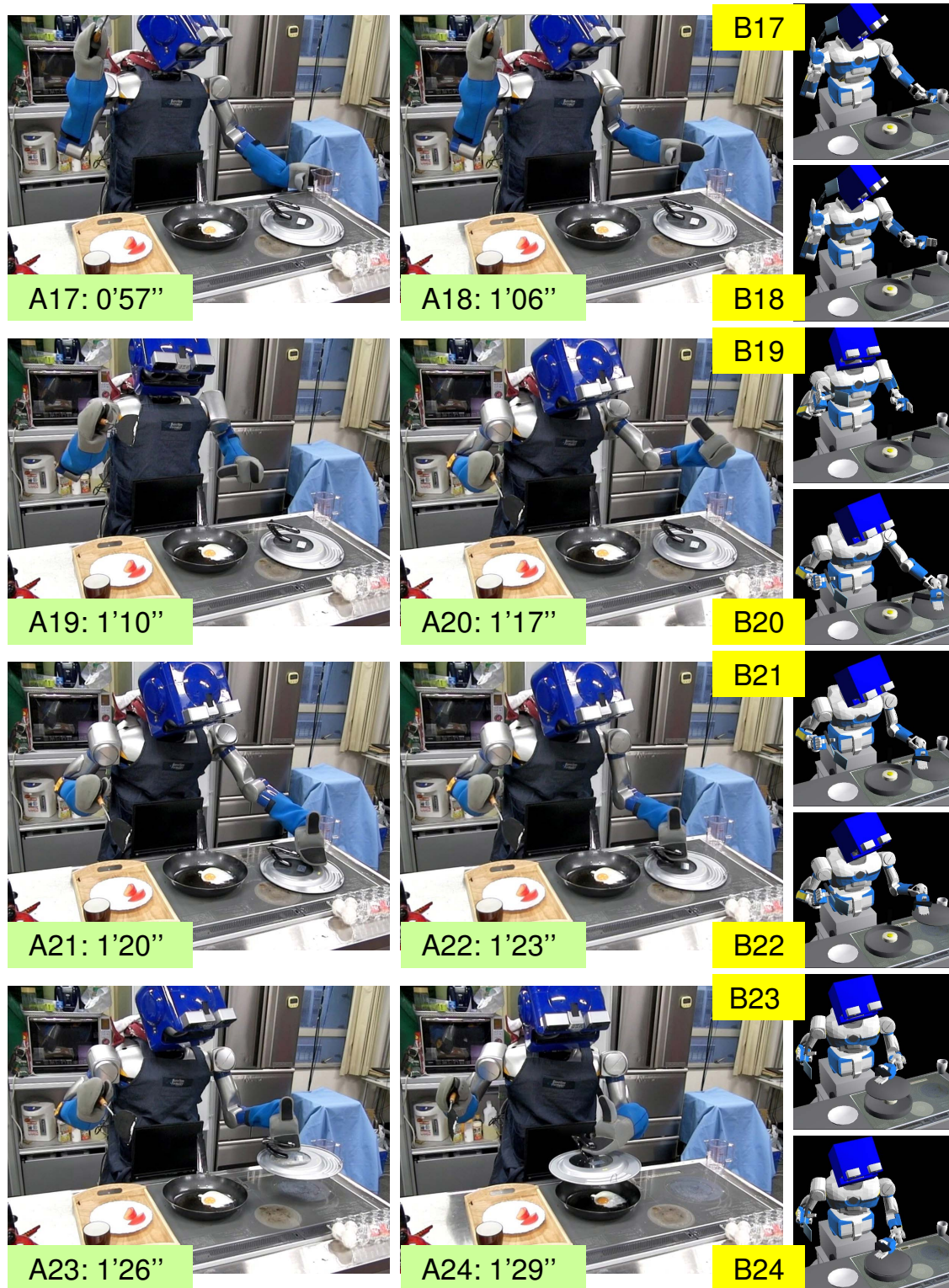


図 3.14: HRP-2VZ による目玉焼き調理実験 (3/6)

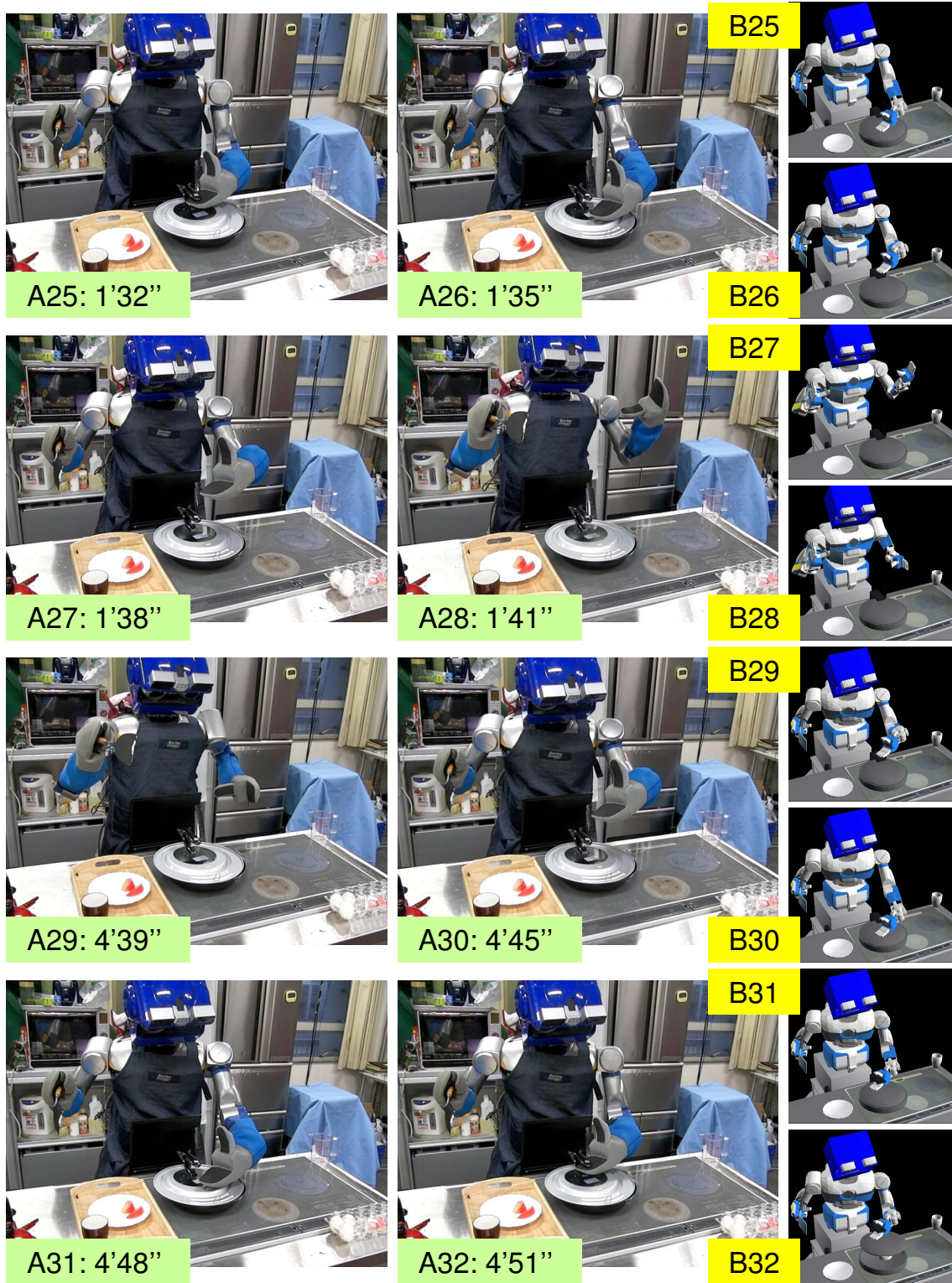


図 3.15: HRP-2VZ による目玉焼き調理実験 (4/6)

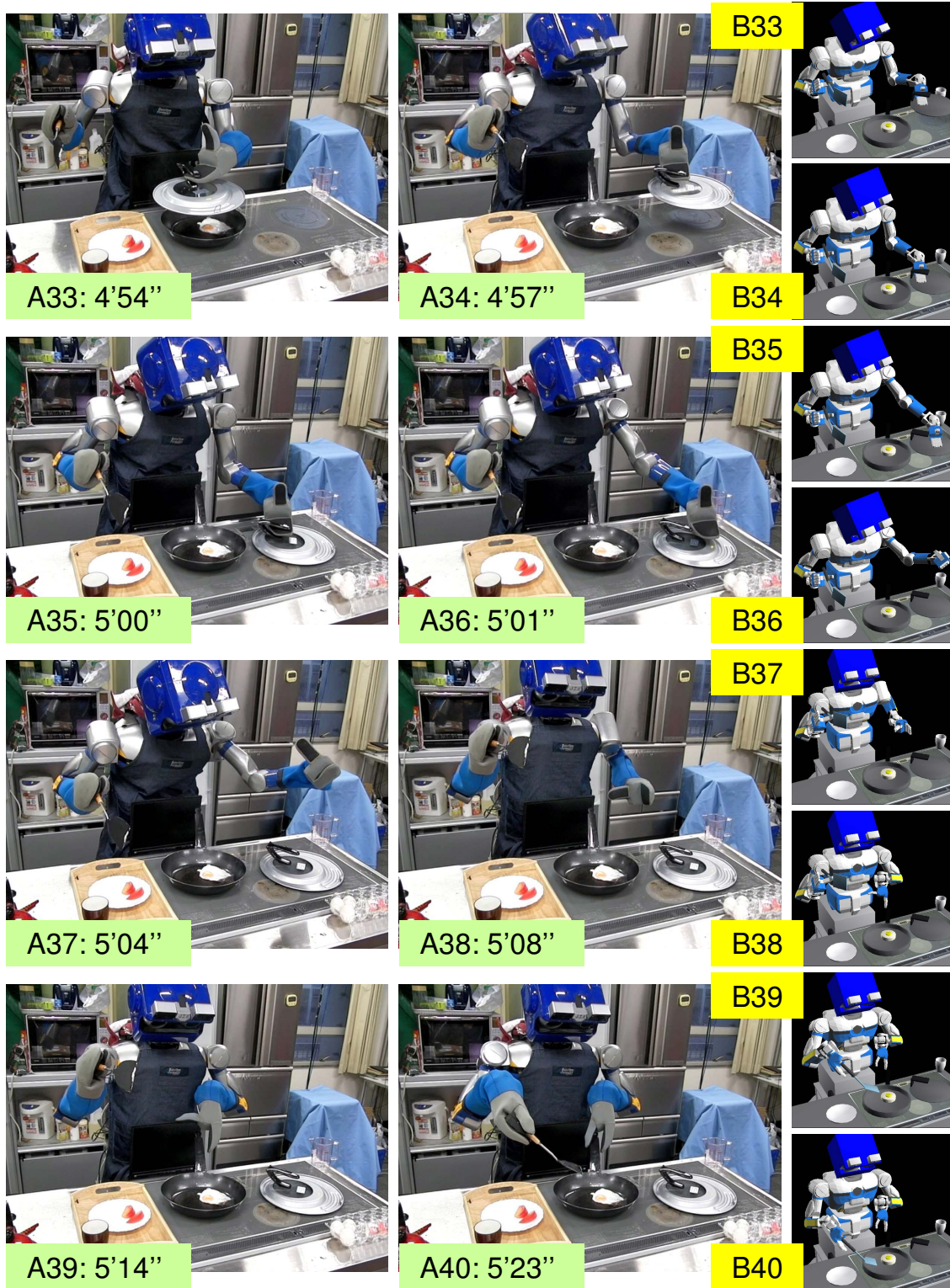


図 3.16: HRP-2VZ による目玉焼き調理実験 (5/6)

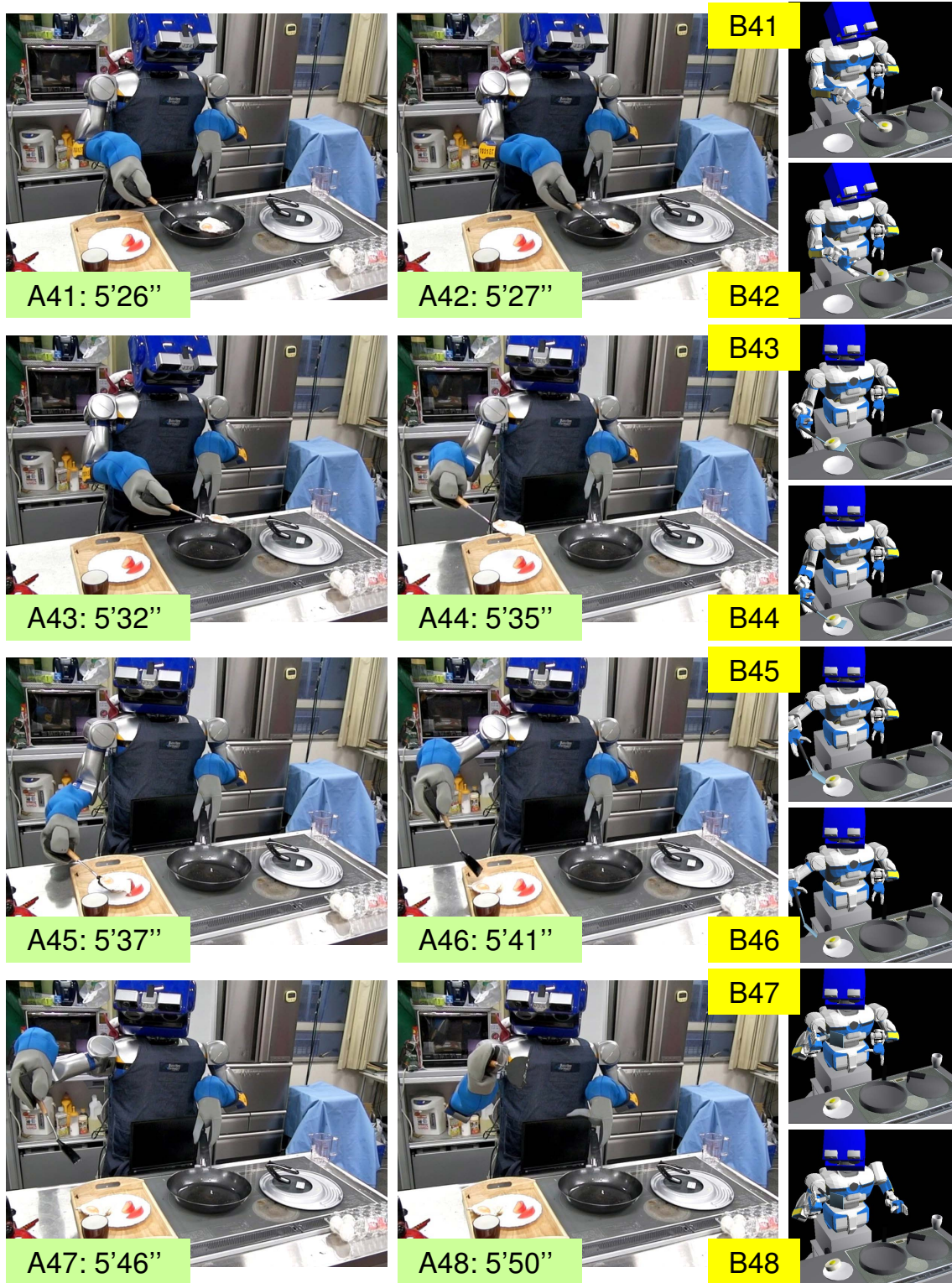


図 3.17: HRP-2VZ による目玉焼き調理実験 (6/6)

3.6 おわりに

本章では、日用道具を用いて生活支援を行うロボットへの要求機能を明らかにし、一つの実現例として、本論文の実験で用いた等身大ヒューマノイド“HRP-2VZ”の構成を示した。HRP-2VZを用い、広視野視覚と高視力視覚を併用し、連続視差フィルタを用いて人間の上半身を追跡しながら、手元へ高視力視覚を誘導する実験では、ロボットが自律的に道具と操作対象を高い解像度で観測することが可能であることを示した。さらに、料理のよそいつけ等の、道具を用いた間接的な物体操作を含む基礎的な調理実験により、道具利用タスクに必要な視覚処理と失敗復帰の重要性について考察した。道具と操作対象の上下・包含・突入関係を確認するための視覚機能と失敗復帰法については、次章で提案し評価する。

第4章

空間遮蔽近傍情報を用いた見えない道具の状態推定

4.1 はじめに

本章では、生活支援ロボットが人を観察し、人から作業を学ぶ際（図 4.1 ）、あるいはロボット自身がタスクをおこなう際の、道具と操作対象あるいは複数の操作対象の間に表れる作用の視覚推定法について述べる。

本章で扱うのは、特に道具と操作対象あるいは複数の操作対象が近接あるいは接触し、強い遮蔽関係が生じるものの一部分は観測可能である状況である。このような状況でも、遮蔽されている領域の近傍の情報（空間遮蔽近傍情報）を用いて双方を追跡し続け、視覚重畳関係を計算し、運動の随伴性という視覚的手がかりを併用することで、上下・突入と部分的な包含関係の変化や、それに伴う相互作用の推定する手法を提案するとともに、本手法を用いたロボット自身の作業の失敗復帰あるいは動作学習についても述べる。

従って本章で検証するのは、2章で述べた以下の事項である。

1. 日常的な道具と操作対象の間の遮蔽関係の観測の可能性
2. 視覚重畳関係を用いた物体同士の関係推定の可能性
3. 視覚重畳関係を用いた物体同士の関係推定を用いた動作学習の可能性
4. 生活支援タスクの学習・実行に際しての有用性

具体例として取り上げるのは、包丁、フライ返しのような平面形状を有する道具が、野菜のような凸形状の物体に作用する場面、あるいはファイルのような凹形状を有する対象に対象の大部分が入り込むような場面である。前者においては上下・突入関係の変化が、後者においては上下と部分的な包含関係の変化が起こり得る。さらに、このような日用道具と操作対象にはテクスチャが少ないという問題点が存在する。

遮蔽された対象を追跡する先行研究としては、局所相関演算を行うためのテンプレート分割して、遮蔽状態の発見および遮蔽下の追跡を行う手法 [37]、対象上または対象周辺の局所特徴を組み合わせて追跡を行う手法 [30] などが挙げられる。また、ある程度遮蔽が生じて、遮蔽されていない部位の局所特徴から物体探索を実

現している研究 [46] がある。しかしこれらの手法では、追跡対象にテクスチャが十分存在している必要があった。

本章で提案する手法では、道具と操作対象あるいは複数の操作対象（以下、追跡対象と呼ぶ）の領域を、三次元空間の中で移動・回転する平面領域として近似し追跡する。このとき、領域の境界位置と内部の色情報を組み合わせた追跡を行う。このような情報の組み合わせによる領域追跡法としては、文献 [63] が挙げられるが、大幅な遮蔽へ対応する枠組みではない。そこで、それらの領域の重畳関係を計算して追跡対象のテンプレートを随時更新することで、大幅な遮蔽へ対応するものとする。さらに、計算された重畳関係と双方の運動の随伴性を併用することで、上下・包含・突入関係の変化を含む作用の分類をおこなうものとする。追跡対象の重なり関係をモデリングしている研究としては文献 [119] が挙げられるが、本研究の手法とは異なり追跡対象の領域を特定していないため、重畳領域や相互重畳関係の計算は不可能である。

従って、本章で提案する追跡法そのものは、操作対象や道具が平面近似可能、つまり回転に起因する見た目の変化が小さいことを前提条件とする。しかし本章で提案する重畳関係の推定法およびその活用法に関しては、2章で考察した通り、二つの立体同士の関係について一般に成り立つものである。追跡時の尤度計算変更方法を変更することで、一般の対象に関しても同様の計算をおこなうことができると考えられる。

なお本章の記述は、著者論文 [53, 54, 99, 101, 102, 106] の内容を含む。

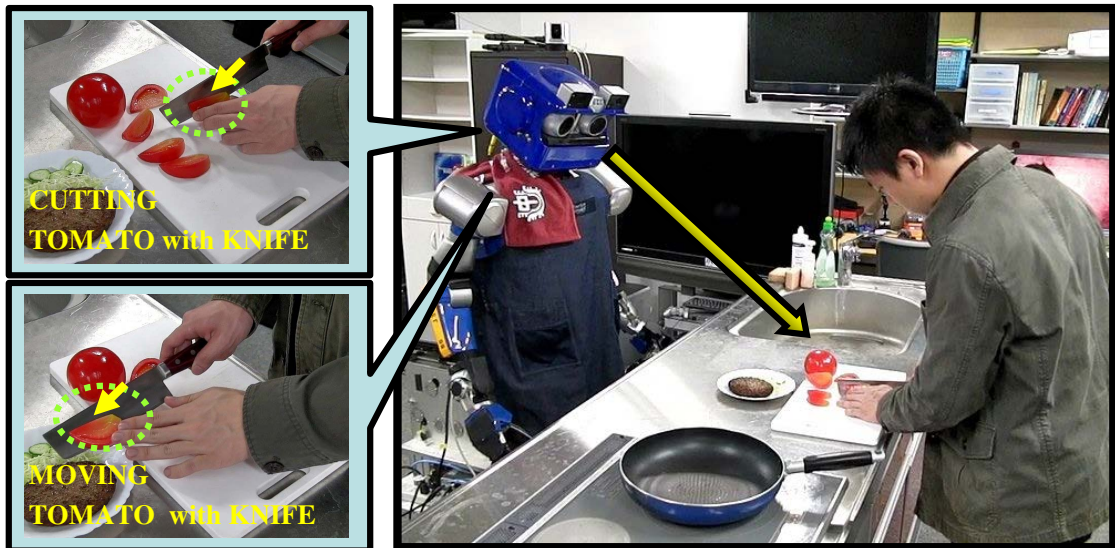


図 4.1: 同じ道具に複数の機能が存在する例と，ロボットによる人の動作観察状況（著者論文 [53] より）

4.2 一部が見えない道具や操作対象の状態推定のための視覚機能

4.3 遮蔽関係のある複数平面領域の追跡法

本節では，テクスチャの少ない平面状の追跡対象領域を，遮蔽下で追跡するための手法について述べる．追跡対象は，初期時刻 $t = 0$ における視覚画像内の領域 $R(0) \subset \mathbb{N}^2$ と座標 ${}^C H_0$ が指定されているものとする．

4.3.1 粒子フィルタを用いた平面道具追跡

本手法では，粒子フィルタを用いて対象を追跡する．粒子フィルタの状態変数 \mathbf{x}_t は，対象のカメラ座標系での三次元位置とロール・ピッチ・ヨー角度，

$\mathbf{x}_t = [x_t \ y_t \ z_t \ \alpha_t \ \beta_t \ \gamma_t]^T$ とする．

追跡対象の初期座標 \mathbf{x}_0 が設定されると，初期粒子 $s_0^{(k)}$ が初期座標へ $s_0^{(k)} = \{\mathbf{x}_0, w_0\}$ のように，一様重み w_0 で設定される．ただし， k は粒子のインデックスである．動作モデル $p(X_t|X_{t-1})$ を $\mathbf{x}_{t|t-1}^{(k)} = \mathbf{x}_{t-1}^{(k)}$ として遷移させた後に尤度計算を行い，尤度 $w_t^{(k)}$ に応じてリサンプリングし，以後それを繰り返す．

4.3.2 尤度計算法

尤度計算の際には，初期座標 ${}^C H_0$ ，初期時刻の対象領域 $R(0)$ および対象画像 $I_A(0) \subset \mathbb{N}^2 \times [0, 255]^3$ から，状態 $\mathbf{x}_t^{(k)}$ の場合の対象領域 $R(t)^{(k)} \subset \mathbb{N}^2$ および対象画像 $I_A(t)^{(k)} \subset \mathbb{N}^2 \times [0, 255]^3$ を生成する．ただし $I_A(0)$ は，初期時刻における入力画像 $I(0)$ のうち $R(0)$ 領域内のみを表すものである．その後，入力画像 $I(t)$ と対象画像 $I_A(t)^{(k)}$ の類似度として尤度を計算する．

また，尤度は以下のような，1) 境界情報に基づく評価，2) 領域内部の色情報に基づく評価，の二種類を統合して求めることで，テクスチャの少ない追跡対象へ対応するものとする．

境界情報に基づく評価

境界情報に基づく評価は，chamfer distance[27]に基づいて行う．この計算では，まず時刻 t の入力画像 $I(t)$ から canny 法 [18] によりエッジを抽出する．そして，画像中の各画素について，その画素から最近傍のエッジまでの距離を計算した，距離変換画像を生成する．その後，対象画像 $I_A(t)^{(k)}$ の輪郭と距離変換画像を用い，chamfer distance を，式 (4.1) により計算する．

$$D_{ch}(P, I) = \frac{1}{|P|} \sum_{p \in P} d_{I(t)}(p) \quad (4.1)$$

ただし， $P \subset \mathbb{N}^2$ は，対象画像の輪郭ピクセルの集合， $|P|$ は P の要素数， $d_{I(t)}(p)$ は，距離変換画像の位置 p における値を表すものとする．従って，変形後の対象画像 $I_A(t)^{(k)}$ の輪郭が入力画像 $I(t)$ のエッジに類似するほど， $D_{ch}(P, I)$ の値は小さくなる．

境界情報に基づく評価値は， $D_{ch}(P, I)$ を用い，式 (4.2) のように求める．ただし， G_e および C_e は定数とする．

$$E_e(\mathbf{x}_t^{(k)}) = \frac{1}{1 + \exp\{-G_e \cdot (-D_{ch}(P(t)^{(k)}, I(t)) + C_e)\}} \quad (4.2)$$

領域内部の色情報に基づく評価

領域の境界位置に基づく評価に加え，領域内部の色を用いた評価 (4.3) を用いる．これは，追跡対象の輪郭の近傍に，他の対象の輪郭が存在した場合に対応するためである．対象領域 $R(t)^{(k)}$ の輪郭周辺について，入力画像 $I(t)$ から対象色の画素を抽出した画素集合 $R_c(t) \subset \mathbb{N}^2$ と， $R(t)^{(k)}$ 内部で輪郭周辺にある画素集合 $R_e(t)^{(k)} \subset \mathbb{N}^2$ の差集合を用いて，式 (4.3) のように求める．

$$E_c(\mathbf{x}_t^{(k)}) = 1 - \frac{|(R_c(t) - R_e(t)^{(k)}) \cup (R_e(t)^{(k)} - R_c(t))|}{|R_e(t)^{(k)}|} \quad (4.3)$$

ただし，記号 $-$ は差集合， $||$ は画素集合の要素数を表すものとする． $E_c(\mathbf{x}_t^{(k)})$ は，0 から 1 までの値をとり，色領域同士の差分が小さくなるほど，評価が高くなるよう設計されている．

評価の統合

二種類の評価値を，式 (4.4) を用いて統合する．

$$E(\mathbf{x}_t^{(k)}) = E_c(\mathbf{x}_t^{(k)})^{C_c} \cdot E_e(\mathbf{x}_t^{(k)}) \quad (4.4)$$

ただし， C_c は正の定数である． $E(\mathbf{x}_t^{(k)})$ は 0 から 1 までの値をとり，評価が高いほど 1 に近づくよう設計されている．

4.3.3 非遮蔽予測領域を用いた対象領域の追跡法

前述の追跡処理を，遮蔽関係を考慮した追跡法へ拡張する．追跡対象が大幅な遮蔽下にある場合，対象の全ての領域を用いた尤度計算をおこなうと，見えていない部分の特徴と類似した，対象上ではない箇所へと誤追跡してしまう可能性が高い．この点に関しては後の基礎実験で検証をおこなう．

そこで本手法では，追跡対象のうち，次の時刻でも見えているであろう非遮蔽予測領域のみの特徴を用いた尤度計算をおこなうものとした．つまり，時刻 T [frame] において遮蔽されていた部分を除いた領域（非遮蔽領域）のみを用いて，時刻 $T+1$ [frame] の尤度計算を行う（図 4.4 の右）．従って，各時刻において，各々の追跡対象の領域内にあって，他の追跡対象に遮蔽されている領域を計算する必要があるが，この計算法については，重畳関係の計算法として次節で述べる．

4.3.4 非遮蔽予測領域を用いた複数対象の追跡・基礎実験

本節で述べた，非遮蔽予測領域を用いた対象の追跡法の有用性を評価するために，平面状の簡単な二つの図形を動かしている状況を撮影し，それらの図形を追跡する基礎実験をおこなった．実験結果を図 4.2 に示す．実験では，人が赤い四角形のプレートと青い円形のプレートを持ち，木の平面板の上で同時に動かす．途中，赤い四角形プレートが青い円形プレートの上に乗る時刻，あるいは青い円形プレートが赤い四角形プレートの上に乗る時刻が存在し，強い遮蔽関係が発生する．

実験では，通常のパーティクルフィルタを用いた追跡器との比較をおこなうため，提案手法での追跡と同時に，テンプレートを固定した，つまり全ての時刻で全ての

対象領域を用いて尤度計算をおこなう手法で追跡をおこなった。尤度計算法は本節で述べた手法を用いるが、提案手法に対して従来法と呼ぶ。ただし、いずれの実験においても、追跡対象は木の平面上のみを移動するものとし、パーティクルフィルタの状態変数は二次元、追跡は木の平面に沿うような並進方向 (x, y) のみでおこなうものとした。

結果、従来法 (図 4.2 の B) では、はじめに赤い四角形プレートが青い円形プレートに覆いかぶさった時刻では正しく追跡ができていたものの、青い円形プレートが赤い四角形プレートに覆いかぶさった時刻では、追跡に失敗した。前者で成功した理由は、青い円形プレートの一部が遮蔽下にあっても、実際の円形プレートの位置で一番尤度が高くなっていたことにあり、これはパーティクルフィルタの、多少の遮蔽下でも頑強な追跡が可能であるという特性によるものだと考えられる。しかし、後者では赤い四角形のほぼ全ての領域が遮蔽されており、尤度のピークが実際の赤い四角形プレートの場所に来なかったことが追跡失敗の原因だと考えられる。一方で、提案手法 (図 4.2 の A) では、いずれの遮蔽状況においても、二つの追跡対象を正しく捉え続けられた。これは、非遮蔽予測領域のみを用い、テンプレートを毎時刻更新し続けることで、テンプレートが小さくなった場合においても、周辺の類似した領域よりも実際の対象領域のほうが尤度が高くなったからだと考えられる。

この基礎実験により、重畳関係を考慮し、非遮蔽予測領域のみを用いた追跡が、重なり合う二つの観察対象の追跡に有用であることが示された。

4.4 重畳関係と随伴性の評価法と作用推定

本節では、遮蔽下にある道具と操作対象あるいは複数の操作対象 (本節では便宜上、一方を道具、もう一方を操作対象と呼ぶものとする) の領域追跡結果から、重畳関係と随伴性を評価して相互作用を推定する手法について述べる。本節の手法により推定された重畳関係は前述の通り、遮蔽下の対象領域追跡における尤度計算にも用いられる。

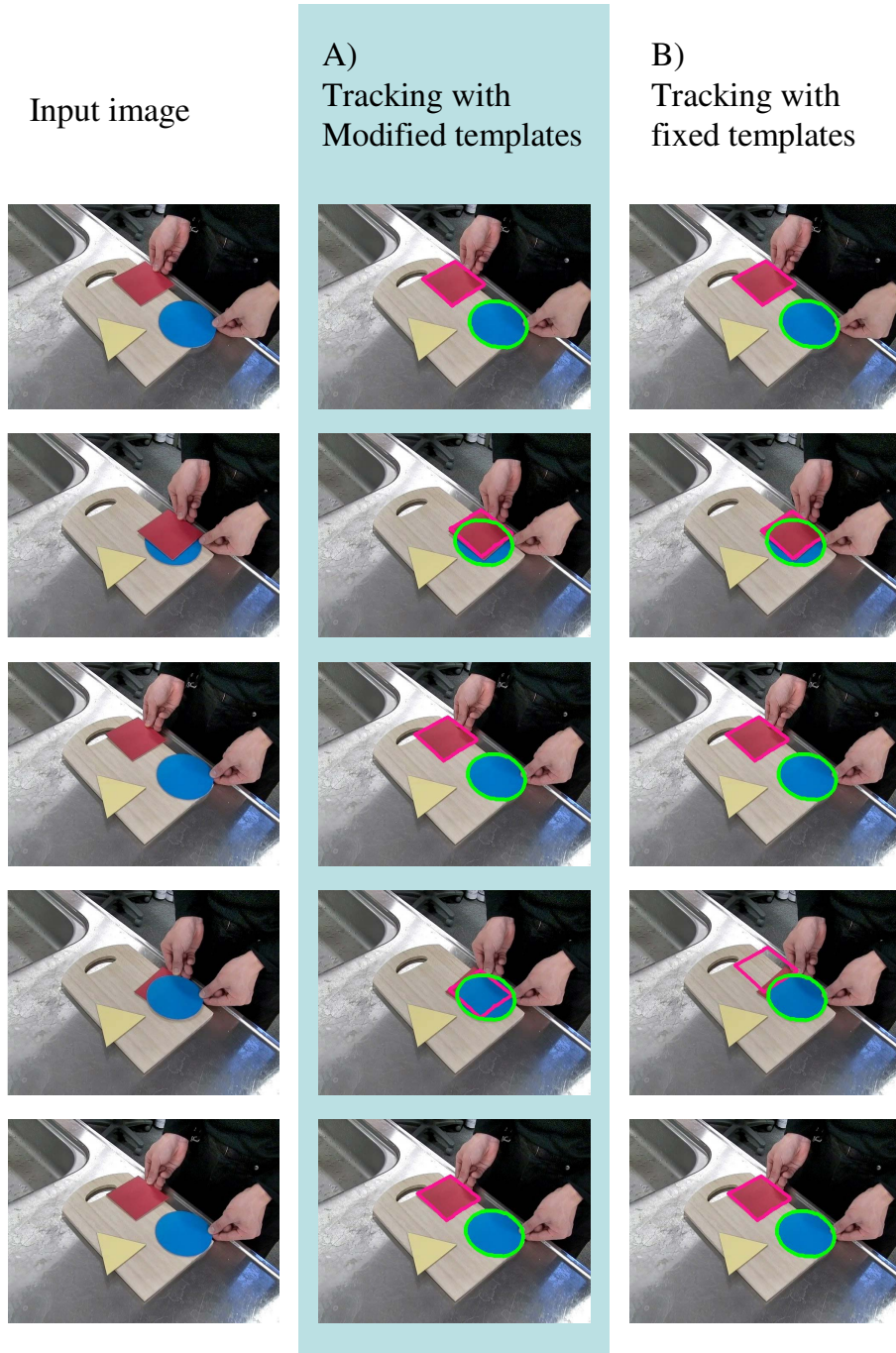


図 4.2: 非遮蔽予測領域のみを用いた追跡の基礎実験：重なりあう二つの板状物体の追跡結果．A) が提案手法，B) はテンプレートを固定して追跡した結果である（著者論文 [53] の図を改変）

4.4.1 作用推定処理の流れ

随伴性と重畳関係を用いて、道具と操作対象の相互作用を推定する処理の流れを図4.3に示す。初期時刻における、道具と操作対象の座標および領域が指定されると、カメラ画像を入力として、双方の追跡が並列に行われる。各々の時刻において、追跡結果となる道具と操作対象の領域画像 $I_A(t)$, $I_B(t)$ は重畳関係の評価器へ通知される。同時に、道具と操作対象の速度 $v_A(t)$, $v_B(t)$ が随伴性の評価器へ渡される。

重畳関係の推定器は、道具と操作対象の領域画像 $I_A(t)$, $I_B(t)$ から重畳関係 $S_s(t) \in \{\text{“Independent”}, \text{“Covered (One-way と同義)”}, \text{“Mutual”}\}$ を計算し、作用の推定器へと通知する。同時に、追跡器へ遮蔽領域が通知され、非遮蔽領域を用いた尤度計算へと利用される。随伴性の推定器は、道具と操作対象の速度 $v_A(t)$, $v_B(t)$ から随伴性 $S_a(t) \in \{\text{“Unclear”}, \text{“No”}, \text{“Yes”}\}$ を計算して、作用の推定器へと通知する。

作用の推定器では、推定された重畳関係と運動の随伴性を用いて、各時刻における作用の種類が判定される。2章での議論によると、相互重畳関係はそれ自体が突入関係あるいは一部包含の関係を推定するのに十分な視覚の手がかりであった。一方で、片側遮蔽関係から上下関係、つまり道具が対象を支えているような関係を推定するには、道具と操作対象が接触しているか否かを検証する必要がある。そこで本章では、道具と操作対象に運動の随伴性が生じていることを、双方が接触している根拠として用いることで、支える関係を推定するものとする。従って、推定できる作用は以下の通りである。

Cut (/Insert)

道具が対象へ突入する作用である。例として対象への変形を含む挿入、切断等が挙げられる。凹状の対象が含まれるならば、少なくともその凸包同士は交わり関係にあることから、一部包含の作用となる。

Move

道具が対象を支えることで移動させる作用。例として、運搬、寄せる等が挙げられる。

Fork

道具が対象へ突入しながら、対象を動かす作用である。例として、フォークや

串で刺して移動する等が挙げられる。

nil

作用が発生していないか、作用の種類を特定できない。

作用の分類は表 4.1 に従っておこなわれ、最終的に、作用の推定結果 $S_f(t) \in \{\text{“Cut”}, \text{“Move”}, \text{“Fork”}, \text{“nil”}\}$ が出力される。

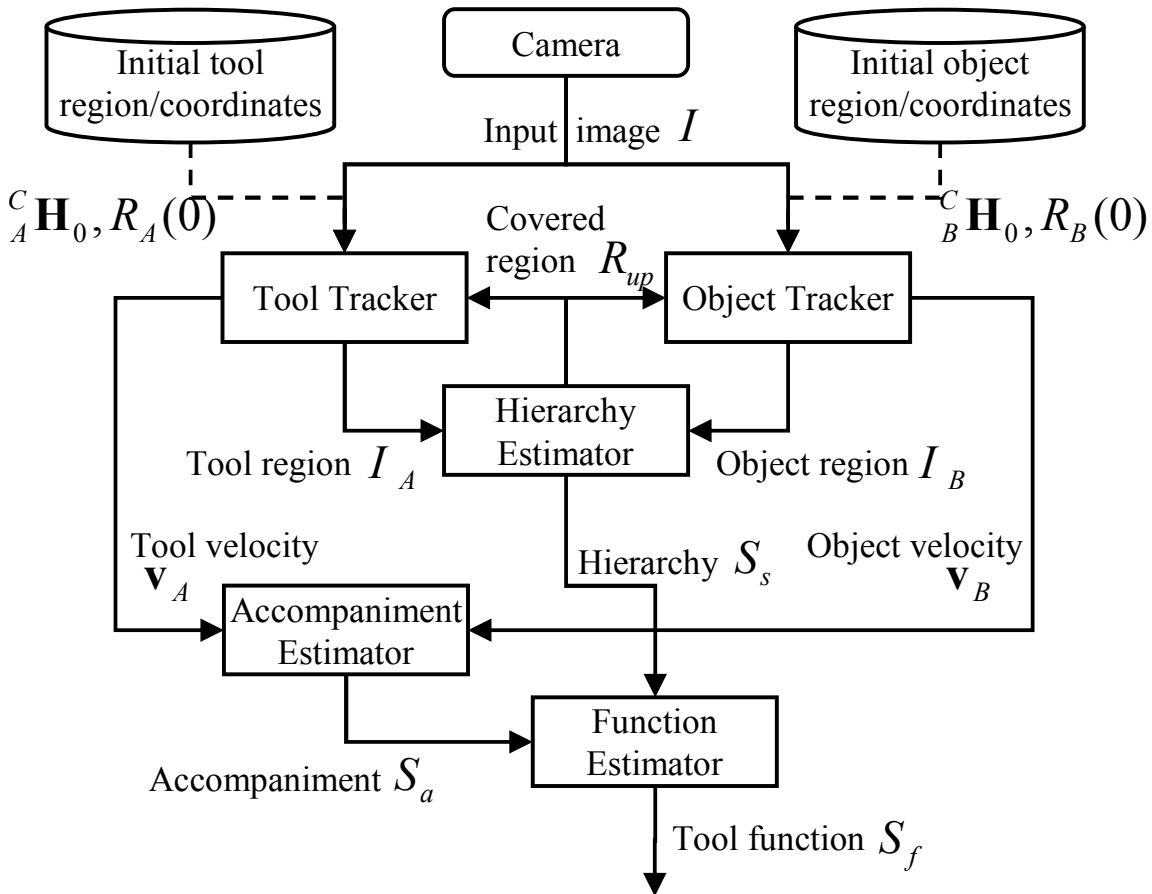


図 4.3: 一部分を遮蔽し合う道具と操作対象を追跡し、それらの間の作用を推定するための処理の流れ (著者論文 [102] より)

表 4.1: 運動随伴性と視覚重畳関係を併用した作用推定

		Superposition		
		Independent	Covered	Mutual
Accompaniment	Unclear	nil	nil	Cut
	No	nil	nil	Cut
	Yes	Move, if contiguous	Move	Fork

4.4.2 重畳関係の分類法

重畳関係の評価器は，時刻 t [frame] において，追跡対象領域の追跡結果 $I_A(t), I_B(t) \subset \mathbb{N}^2 \times [0, 255]^3$ が通知されると，対象領域 $R_A(t), R_B(t) \subset \mathbb{N}^2$ 同士の重なり領域 $R_{A,B}(t) = R_A(t) \cap R_B(t) \subset \mathbb{N}^2$ を求める（図 4.4 の点線領域）．その後， $R_{A,B}$ 内部の各画素について，入力画像 $I(t)$ が各対象画像 $I_A(t), I_B(t)$ のいずれに近いかを計算し，ラベリングする（図 4.4 左の矢印）．対象同士の重畳関係を決定する研究としては，物体の持ち込みあるいは持ち出しの観測中に，物体同士の重なり関係の発生あるいは消失を扱った研究 [47] がある．しかし，本研究では物体への突入関係も扱えるように，相互重畳関係をも推定できる手法を導入する．

ラベリングおよび重畳関係の決定法を，式 (4.5) に示す．ただし， $I(p) \in [0, 255]^3$ は，入力画像 I の位置 p における輝度値 (R,G,B) とし，記号 $\|x - y\|$ は，輝度値 x, y 同士の距離を表すものとする．本章の実験では，マンハッタン距離を用いた．また，ノイズ対策のため，記号 \exists と \forall の計算には，閾値処理を行う．

$$S_s(t) = \begin{cases} \text{“Independent”}, & \text{if } R_{A,B} = \emptyset. \\ \text{“Covered”}, & \text{if} \\ & (\forall p \in R_{A,B}; s(p, A, B)) \vee (\forall p \in R_{A,B}; s(p, B, A)). \\ \text{“Mutual”}, & \text{if} \\ & (\exists p \in R_{A,B}; s(p, A, B)) \wedge (\exists p \in R_{A,B}; s(p, B, A)). \end{cases} \quad (4.5)$$

$$s(p, X, Y) \triangleq \|\mathbf{I}(p) - \mathbf{I}_X(p)\| \geq \|\mathbf{I}(p) - \mathbf{I}_Y(p)\|$$

前述の通り，追跡器において，次の時刻 $T + 1$ [frame] における尤度計算は，時刻 T [frame] における非遮蔽領域のみを用いて行う（図 4.4 右）．そのため，このラベ

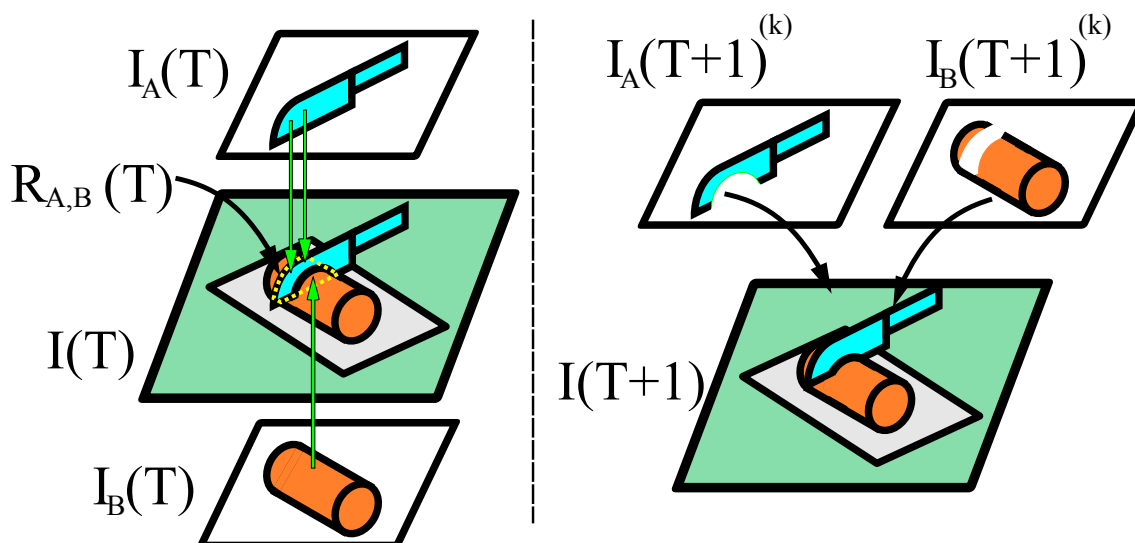


図 4.4: 左：対象領域同士の重畳関係の推定法，右：非遮蔽予測領域を用いた追跡法（著者論文 [53] より）

リング結果は，作用の推定器だけでなく，各々の追跡器へも通知される．

4.4.3 運動随伴性の評価法

運動随伴性認識の先行研究としては，把持物体領域を選択するために，身体運動情報と物体運動情報の正準相関分析を用いる研究 [136] が挙げられる．一方で本追跡システムでは，道具と物体の追跡結果が，いずれも同じ三次元空間内の速度として出力されるため，速度同士の差を直接評価できる．そこで，時刻 t [frame] における道具の速度を $v_A(t)$ ，操作対象の速度を $v_B(t)$ とし，時刻 T [frame] における随伴性を，最新 N [frame] の速度差を平均して，式 (4.6) を用いて評価する．

$$E_{acm}(T) = \exp \left\{ \frac{-1}{C_{acm}} \sum_{t=T-N+1}^T \|v_A(t) - v_B(t)\| \right\} \quad (4.6)$$

ただし， C_{acm} は正の定数とする． $E_{acm}(t)$ は，0 から 1 の間の値をとり，随伴性が強く見られるほど，値は大きくなる．

しかし、静止している物体同士では、随伴性を正しく評価することはできない。そこで、式(4.7)に示す平均速さ $\overline{\|\mathbf{v}_X(t)\|}$ の評価を併用し、随伴性判断の可否を決定する。

$$\overline{\|\mathbf{v}_X(T)\|} = \frac{1}{N} \sum_{t=T-N+1}^T \|\mathbf{v}_X(t)\| \quad (4.7)$$

最終的に、時刻 t における随伴性 $S_a(t) \in \{\text{“Unclear”}, \text{“No”}, \text{“Yes”}\}$ を、式(4.8)のように決定する。つまり、双方ともに動いており、動きに相関がある場合のみ随伴関係があり(“Yes”)、双方ともに停止していた場合は、判断を保留(“Unclear”)、その他の場合は、随伴性が無い(“No”)とする。

$$S_a(t) = \begin{cases} \text{“Yes”}, & \text{if } h(A, t) \wedge h(B, t) \wedge a(t). \\ \text{“No”}, & \text{if } (h(A, t) \wedge h(B, t) \wedge \neg a(t)) \\ & \vee (h(A, t) \wedge \neg h(B, t)) \\ & \vee (\neg h(A, t) \wedge h(B, t)). \\ \text{“Unclear”}, & \text{if } (\neg h(A, t) \wedge \neg h(B, t)). \end{cases} \quad (4.8)$$

$$h(X, t) \triangleq \overline{\|\mathbf{v}_X(t)\|} \geq v_{thr}$$

$$a(t) \triangleq E_{acm}(t) \geq E_{acm.thr}$$

ただし、 v_{thr} および $E_{acm.thr}$ は定数である。

4.5 重畳関係・随伴性を用いた作用分類実験

本節では、前節までに述べた、重畳関係および運動の随伴性の評価法を組み合わせ、人が道具で対象を操作しているシミュレーション画像列、そして実画像列から作用分類をおこなった実験結果について述べる。

4.5.1 作用分類のシミュレーション実験

まず、重畳関係と運動の随伴性の評価を用いる作用推定の提案法が、想定したような作用推定となっているかを確認すべく、シミュレーション実験をおこなった。シミュレーション実験では、euslisp[61, 116] 環境内で操作対象形状モデルと道具形状モデルを動かしている様子を投影した画像を用いる。ただし、対象領域の追跡ミス

に起因する運動の随伴性の誤差がない状態で評価するため、対象領域の追跡はせず、各形状モデルの対象領域 $I_A(t)$, $I_B(t)$ と速度の正解データ $v_A(t)$, $v_B(t)$ を直接用いるものとした。

シミュレーション画像と実験の結果を、図 4.5 および図 4.6 に示す。実験 A, 実験 B では包丁で食品を切る様子を扱ったが、包丁が食品へ切り込んでいるときのみ相互重畳関係が推定され、いずれも“CUT”という作用が推定された。実験 C, 実験 D では包丁で食品を動かす様子を扱った。結果、いずれも後半では相互重畳関係と運動の随伴性から、“MOVE”という作用が推定されている。さらに実験 E では、食品に串を刺して動かす様子を扱った。串を刺す時刻では運動に随伴性が見られないため、対象に切り込む“CUT”という作用が推定され、その後動かした際に“FORK (SKEWER)”という作用が推定された。いずれも、期待した通りの結果が得られている。

4.5.2 実画像を用いた実験の条件

実画像の入力データとしては、人が机上で包丁を使い、果物に作用する場面を撮影した、サイズ 640 [pixel]×480 [pixel], 15 [fps] の画像列を用いて実験をおこなった。実験 1 では果物を切った場面、実験 2 では果物を包丁に載せて運んだ場面を選択した。ただし観測の際には、切断作用が行われる際に相互重畳関係が発生する視点としている。実験 1, 2 のいずれも包丁と果物の領域を追跡し、それらの相互作用を推定するものとした。ただし、包丁と果物は初期時刻において、既知の作業平面上に存在するものとする。初期時刻における追跡対象の領域は、色ヒストグラムを用いたグラフカット法 [11, 137] により、対話的に人が指定するものとした。既知の作業平面座標を用いて、操作対象の初期座標を対象領域の重心位置に設定し、本章で述べた追跡法を用いて両追跡対象の領域を追跡していく。

グラフカット (Graph-cut) は、ユーザがインタラクティブに画像を領域分割するための手法として知られている [11, 137]。また、ロボットがグラフカットにより自律的に領域分割する例としても、著者らによる操作対象物の自律認識研究 [103] のほか、しわ特徴を種として衣類領域分割を行う研究 [76] 等がある。ユーザが画像中の

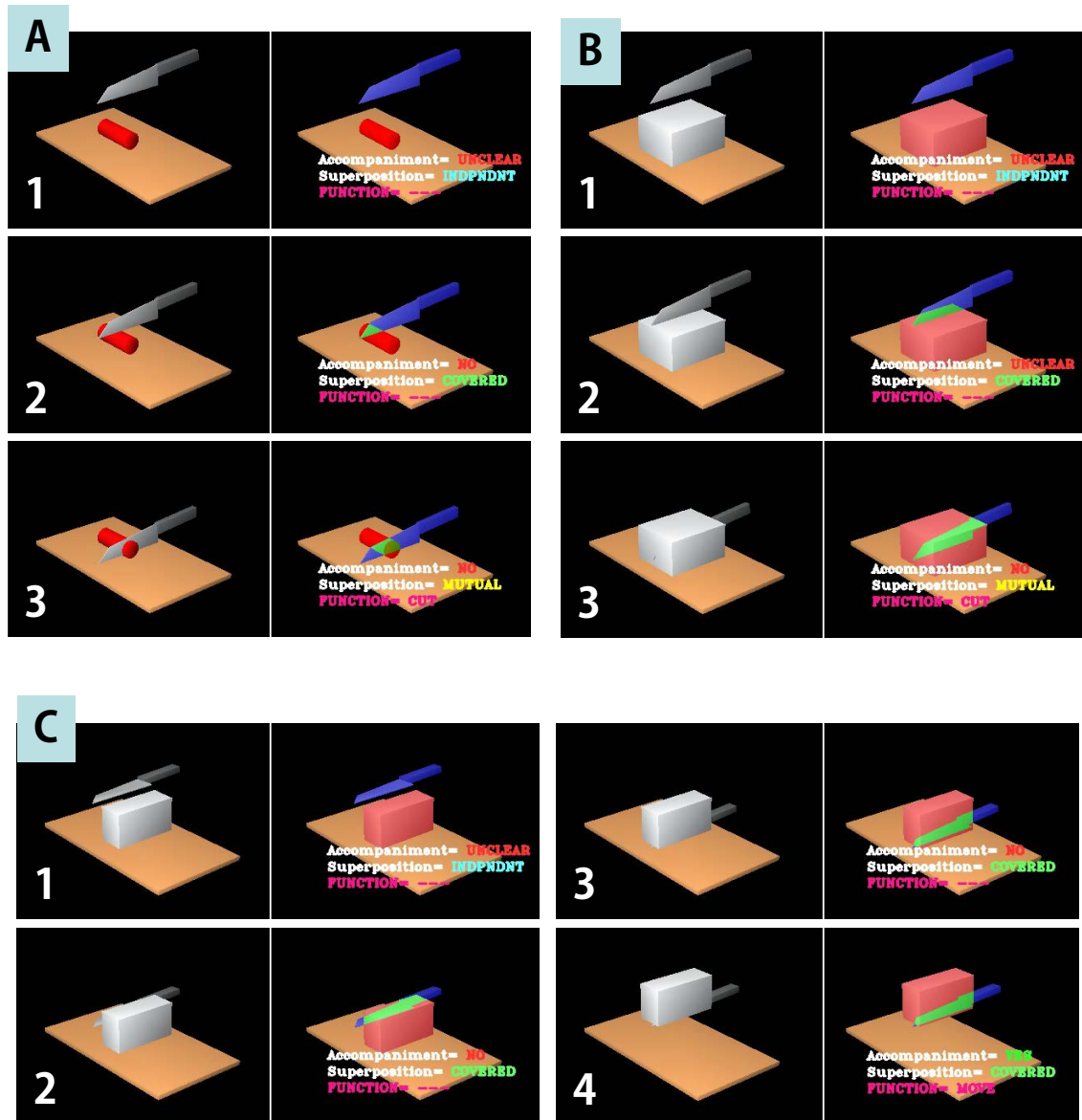


図 4.5: 視覚重畳関係と運動随伴性に基づく作用分類の基礎実験 1

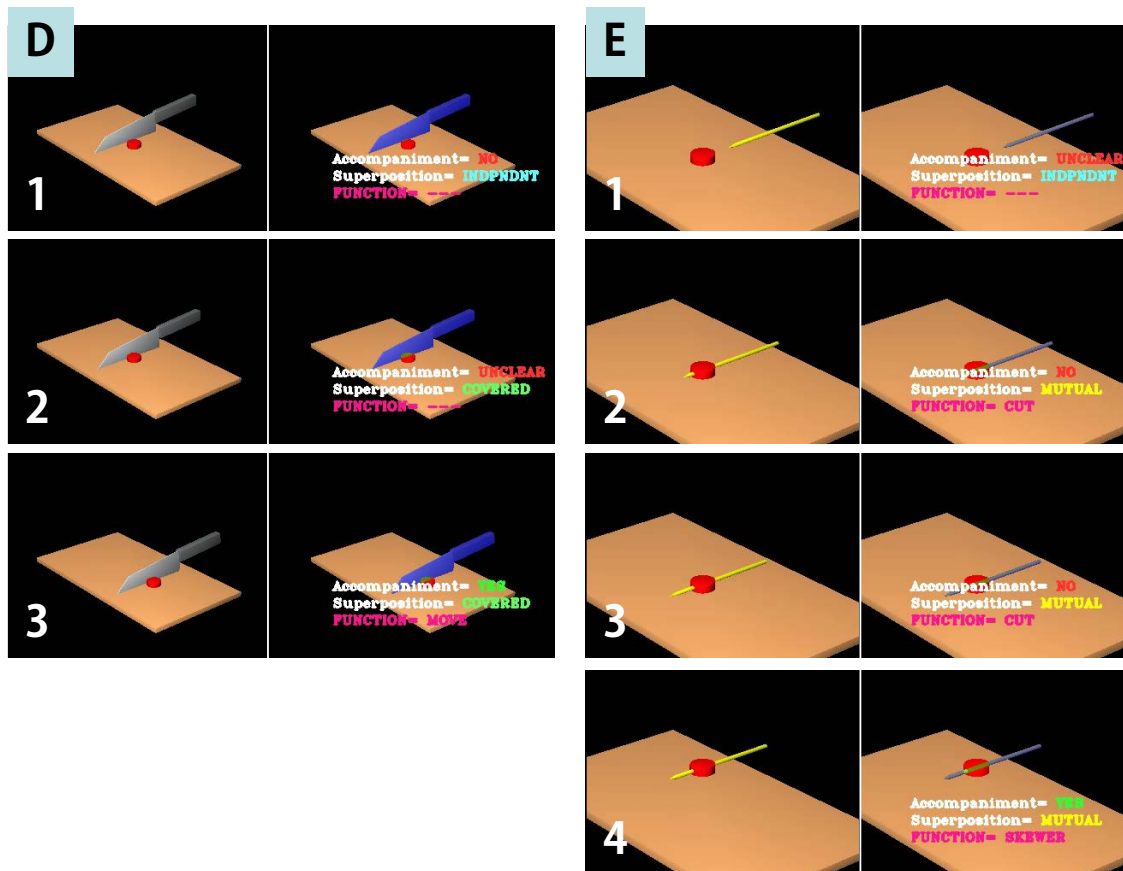


図 4.6: 視覚重畳関係と運動随伴性に基づく作用分類の基礎実験 2. ただし, 作用 “Fork” は “SKEWER” として表示している

前景領域の一部と背景領域の一部をそれぞれ指定すると¹，画像を前景領域と背景領域へ分割するものである．

グラフカットでは，(4.9) に示すコスト関数を最小化するように画像のラベリングをおこなう [137] ．

$$E(X) = \sum_{v \in V} g_v(X_v) + \sum_{(u,v) \in E} h_{uv}(X_u, X_v) \quad (4.9)$$

ただし， V は画素を表す集合， $E \subset V$ は，隣接している画素の集合であり， X_v は画素に割り当てられたラベルを示す．右辺第一項 $g_v(\cdot)$ はデータ項と呼ばれ，各ラベルへ画素が所属する尤度を表す．また，第二項 $h_{uv}(\cdot)$ は平滑化項と呼ばれ，隣接画素間の切断コストを表す．このコスト関数は，背景と似たような色の画素は背景領域へ，前景と似たような色の画素は前景領域へ所属させると同時に，画像内のエッジ上で分割がなされるように設計されたものである．

(4.10) および (4.12) に， $g_v(\cdot)$ ， $h_{uv}(\cdot)$ の計算法を示す．

$$g_{v(p)}(l) = -\log \alpha(I(p), l) \quad (4.10)$$

$$h_{v(p)v(q)}(l, l') = 0 (l = l') \quad (4.11)$$

$$h_{v(p)v(q)}(l, l') = \frac{\text{dist}(p, q)}{\lambda \exp(kI(p) - I(q))^2} (l \neq l') \quad (4.12)$$

ここで $I(p)$ は画素 p の明るさ， l はラベルを示し（1 を前景，2 を背景とした）， $\alpha(I(p), l)$ は明るさ $I(p)$ の頻度を示す．また， λ と k は正の定数である．

(4.9) のコスト関数を最小化するような画像分割の計算は，画像を図 4.7 に示すようなグラフのグラフ切断問題へ帰着することができる．このグラフでは，中段のノード W_u, W_v が各画素を表し，上段ノード t および下段ノード s が，ユーザが指定した前景および背景データを示す．全ての中段ノードには，上段ノードおよび下段ノードからエッジを張る．さらに，中段のノードのうち隣接画素同士にはエッジを張る．ただし，本研究における隣接画素は八近傍画素とする（図 4.7 では簡単のため四近傍のみエッジを張った）．(4.9) のデータ項および平滑化項は各々，前景および背景データと画素ノードの間のエッジ，そして画素間のエッジの切断コストとし

¹前景・背景の一部を描画した“trimap”と呼ばれる画像を作成する

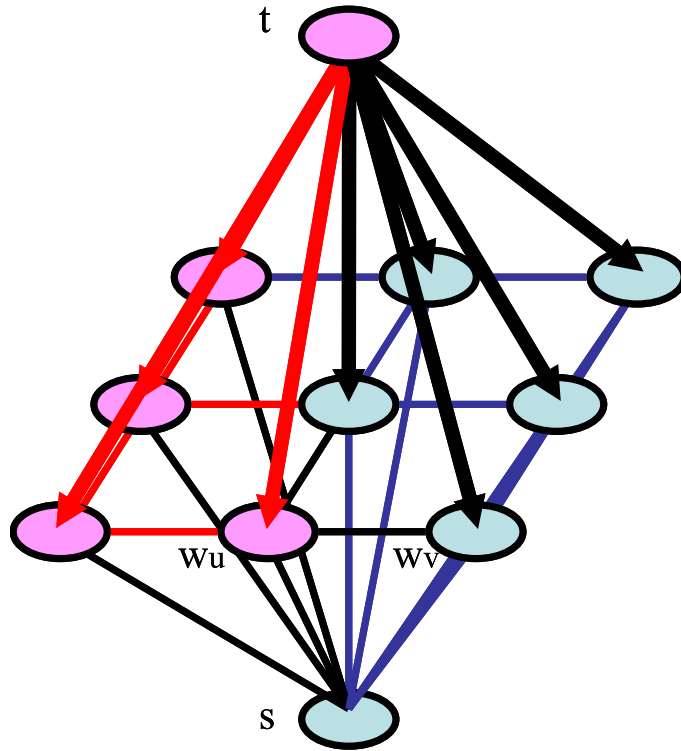


図 4.7: グラフカット法におけるグラフ作成法 (著者論文 [99] より)

て表現する．つまりこのグラフの切断問題は，なるべくデータ項と平滑化項の小さいエッジでグラフを二分する問題となる．この問題は，グラフの最大流問題と同値となり，既存の手法 [11] により解決できる．

本手法を GUI 操作により試験した結果を図 4.8 に示す．A が入力画像，B が trimap (ユーザが作成する，前景・背景指定画像) である．白色が前景指定された領域，黒色が背景指定された領域を示し，C がグラフカット結果である．D1 と D2 は $g_{v(p)}(1)$ ， $g_{v(p)}(2)$ を，E1 から E4 は $h_{v(p)v(q)}(l, l')$ をグレースケールで示したものである．ただし，E1 から E4 はそれぞれ，右，右下，下，左下近傍の画素との評価を示す．

4.5.3 実画像を用いた実験の結果と評価

まず，実験 1 および実験 2 における領域追跡結果を，図 4.9 と図 4.10 に示す．各々の左列が入力画像，右列が領域の追跡結果である．領域の追跡結果は赤色あるいは

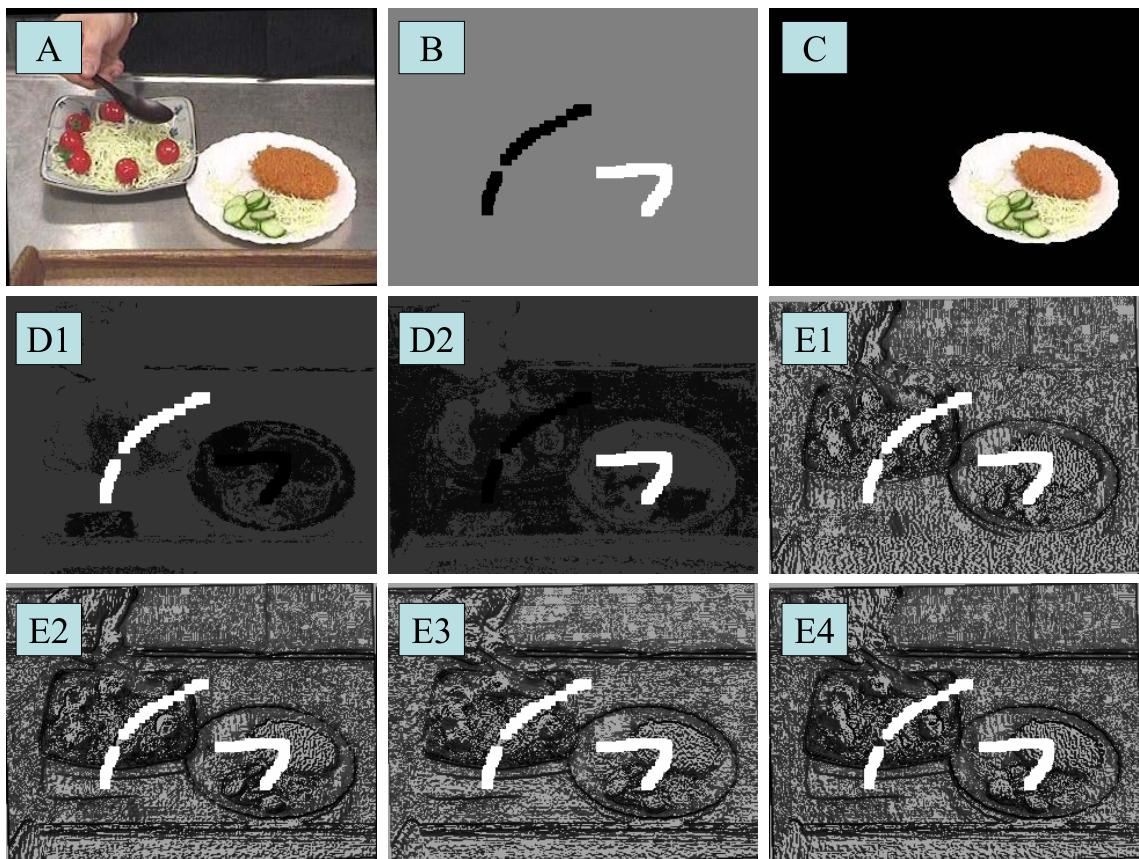


図 4.8: グラフカット法によるインタラクティブな領域分割 (著者論文 [99] より)

緑色の太線で示している。いずれの実験でも、途中で追跡対象同士に遮蔽関係が発生しているが、包丁と果物それぞれの領域追跡に成功している。

次に、実験1および実験2における、各時刻の随伴性と重畳関係および作用推定の結果を図4.11および図4.12に示す。実験1においては、随伴性は無しまたは不明で推移しているが、実験2では80フレーム以降に連続して随伴性有りとして計算されている。重畳関係は、いずれの実験でも最初は独立であったが、実験1では片側遮蔽から相互重畳となり、その後片側遮蔽へと変化した。切断に伴う正しい重畳関係が推定できていると言える。実験2では、後半で片側遮蔽となっているが、相互重畳の時刻も存在する。実際には相互重畳は発生していないが、追跡対象同士に色の近い部分が存在し、上下関係の推定に間違いが生じたことが原因だと考えられる。

実験1において、随伴性と重畳関係を統合した作用の推定結果は、後半でCutとなった。実験2では、後半でForkまたはMoveと推定された。作用推定の成功率は、各々91[%]、65[%]であった。ただし成功率は、正解データとして人が記述した作用Cut、Move、Fork、nilと、同じ作用が推定されたフレーム数を、全てのフレーム数で割って計算した。実験2における成功率の低さは、前述の重畳関係推定の誤りに起因するものだと考えられる。重畳関係推定の誤りは、突入している道具は、その面の法線方向には移動しにくい等、追跡対象同士の重なり方の推移についての常識をモデル化し利用することで改善できると考えられる。また、距離画像を追跡の尤度計算で併用する方法も考えられる。ただし、距離画像においても遮蔽関係は発生するため、この場合も本手法のように重畳関係を陽に考慮する必要がある。

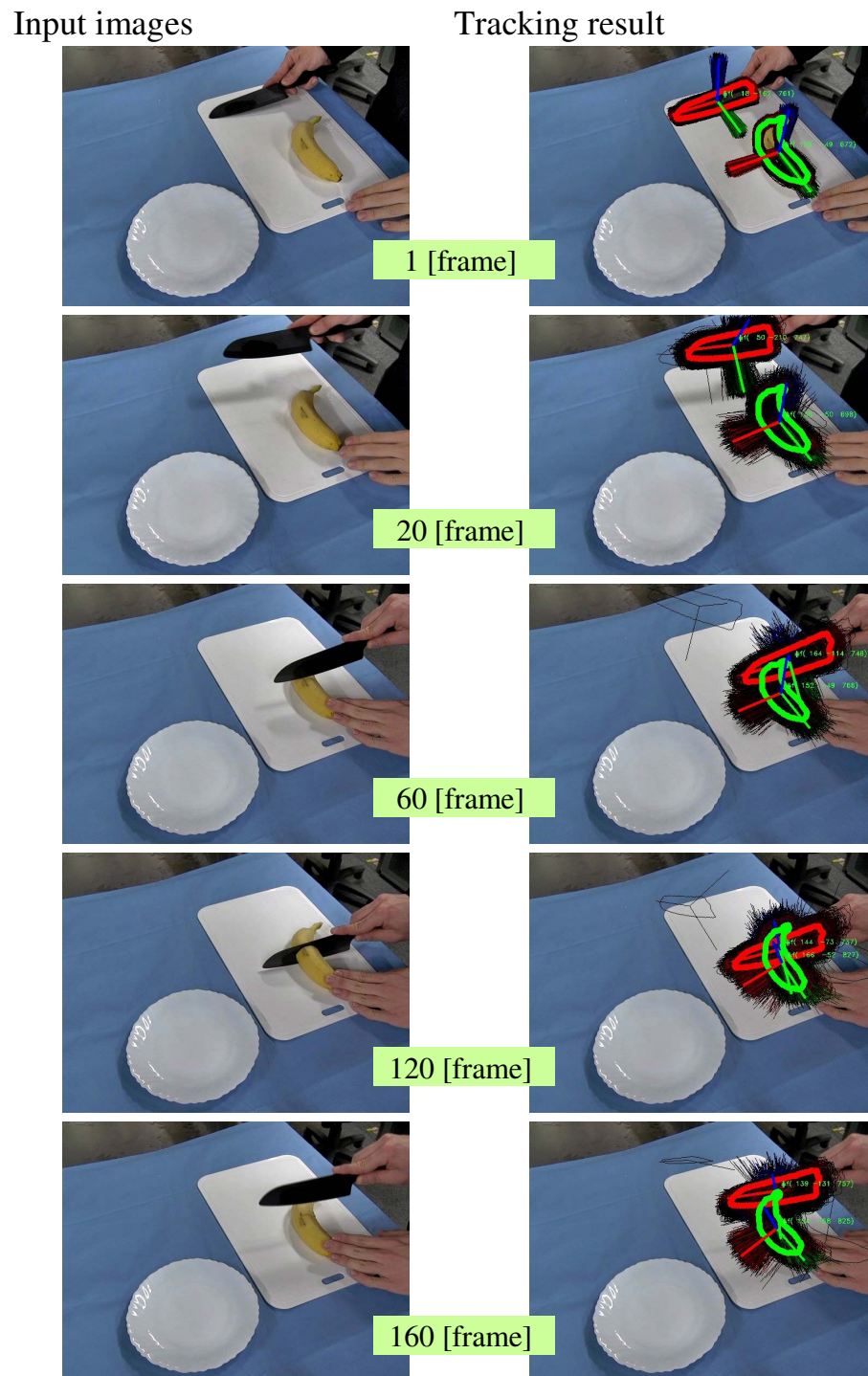


図 4.9: 包丁と果物の追跡実験 1 (切断) の結果 . 赤い枠で示したのが追跡された包丁の領域 , 緑の枠で示したのが果物の領域である . 図中の赤・緑・青の各色の線分は , 包丁と果物の座標系のそれぞれ x , y , z 軸を示す . 黒線で示したのは , 各パーティクルの表現する対象領域と座標系である (著者論文 [53] の図を改変)

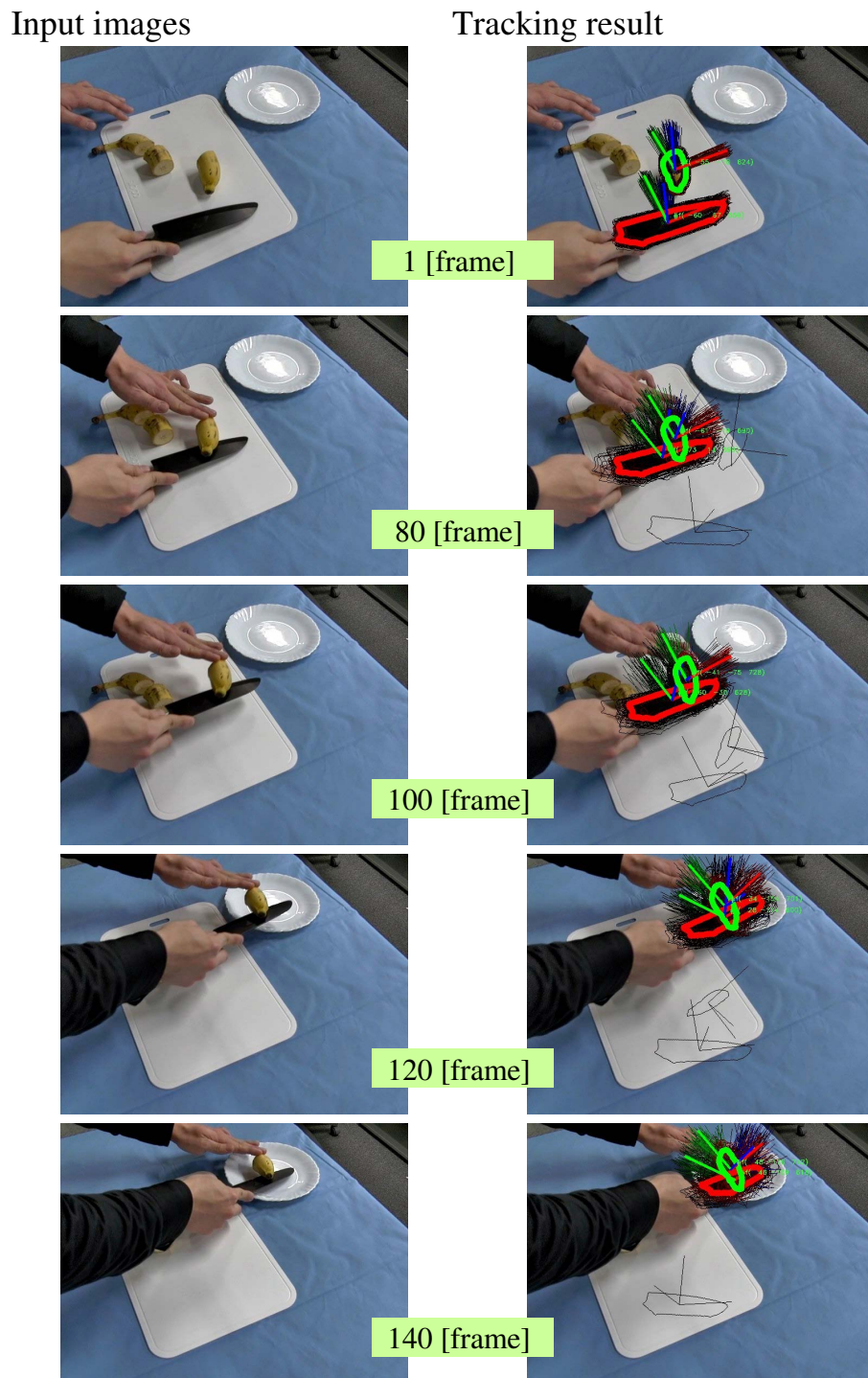


図 4.10: 包丁と果物の追跡実験 2 (移動) の結果．赤い枠で示したのが追跡された包丁の領域，緑の枠で示したのが果物の領域である．図中の赤・緑・青の各色の線分は，包丁と果物の座標系のそれぞれ x , y , z 軸を示す．黒線で示したのは，各パーティクルの表現する対象領域と座標系である（著者論文 [53] の図を改変）

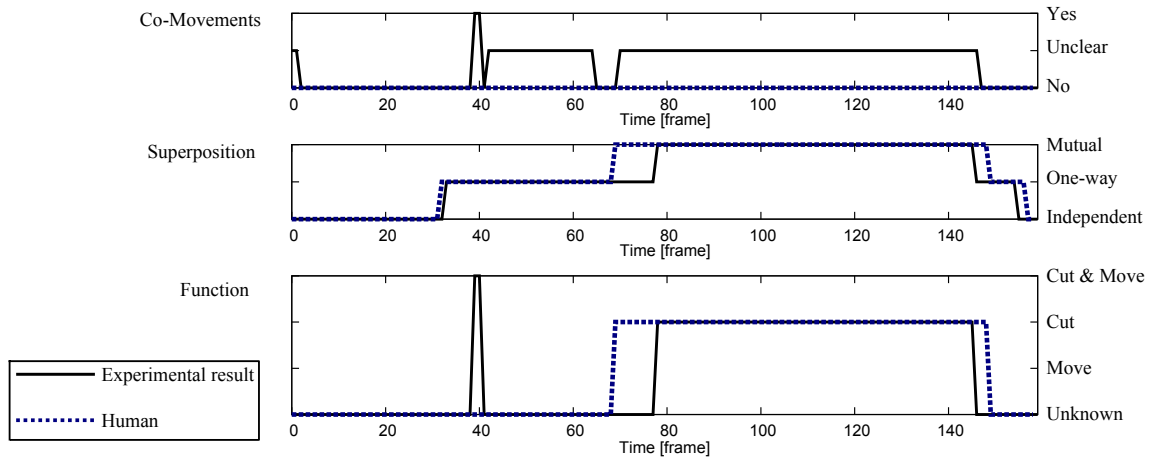


図 4.11: 実験 1 (切断) における運動随伴性, 視覚重畳関係, および作用の推定結果. Unknown は nil と同義である (著者論文 [53] の図を改変)

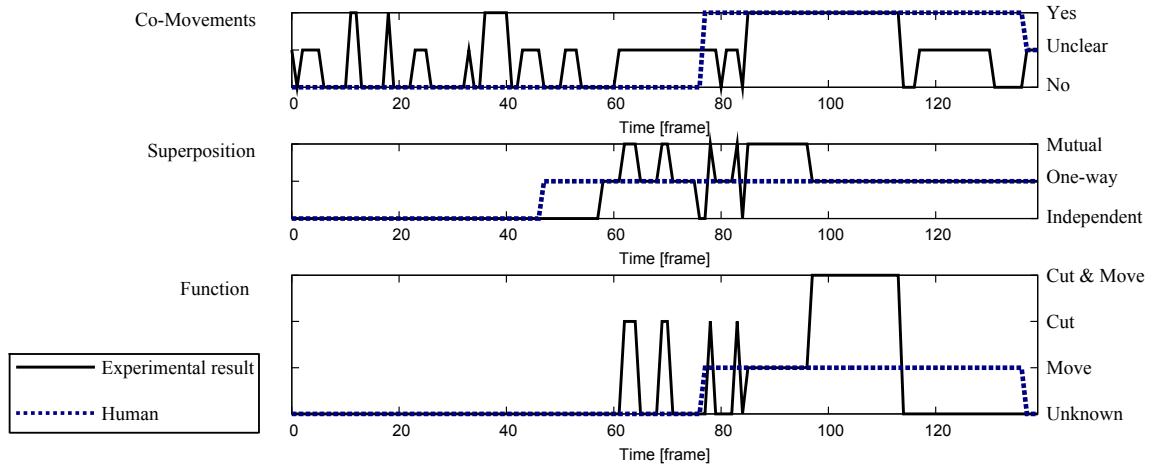


図 4.12: 実験 2 (移動) における運動随伴性, 視覚重畳関係, および作用の推定結果. Unknown は nil と同義である (著者論文 [53] の図を改変)

4.6 ロボットによる見えない対象の状態推定に基づく道具利用

本節では、ロボット自身が道具を利用する状況を想定し、道具と操作対象の上下・突入関係を変化させるような操作を扱う。特に道具や操作対象が薄い場合、あるいは柔軟性を有する場合などでは、操作には失敗が頻繁に起き得る。そこで本節では、本章で提案し人の道具操作観察に用いた、視覚重畳関係を基軸とした物体同士の関係推定法を、ロボット自身が道具を使っている際の道具と操作対象の関係推定に用い、操作の成功・失敗を認識し、認識結果に応じて動作を修正する方法を提案する。この手法では、遮蔽領域の空間近傍の特徴を利用し、物体と物体の重なり方に着目することで物体同士の簡単な物理関係を推定し、この関係を目的状態として操作をやり直させる。この手法にはロボットの視点に制約があるものの、成功・失敗状態表現および失敗からの復帰記述を簡潔に行うことが可能である。さらに実ロボットを用いて

- 1) 柔軟物体同士の挿入操作であるファイリング作業，
 - 2) 薄く摩擦力が変化する対象に対する scooping 操作である，料理のよそいつけ作業
- へ提案手法を導入し，有用性を検証した。

4.6.1 重畳関係を用いた状態推定と動作修正

本章で扱うのは、2章で述べ分類した、道具と操作対象あるいは複数の操作対象（以下、本節では道具と操作対象をまとめて操作対象と呼ぶ）の上下・突入関係を変化させる操作である。各操作について、成功・失敗状態の確認、および修正動作に必要なズレを観測することを考える。このような操作では、互いの広い面を合わせる場合が多い。これは多くの接触関係が重力や摩擦力などにより保たれているからである。しかし、接触面を接触面に近い視点から観測しようとする、その薄さにより観測しにくいという問題がある。さらに、特に重力により重なりあう操作対象同士の接触面は、それらを操作する床面や机上に近い位置となる場合が多く、横から観測するよう視覚センサを配置するのは難しい。

しかし、上下・突入関係を変える操作では、操作対象物が薄い場合、柔軟性を有する場合などにおいて失敗が起こる可能性が高い。そこで、操作をおこなってから操作対象間の状態を認識し、失敗を検知し、失敗状態に応じてやり直しをするようにすれば良いと考える。

この場合においても、上記のような視点に関する制約がある多くの場合では、接触面を接触面の外から（例：上から）観測することとなる。これはデメリットとメリットの両方を有する。デメリットは、多くの場合発生する遮蔽により、接触面に垂直な対象の動きの高精度の観測が難しいことである。しかし、本章で述べてきたような、視覚重畳関係の認識には有利な視点となることがメリットとなる。

このような視点の取り方、視線方向を、図 2.4 の緑線で示した。この緑矢印方向から複数の対象を観測した場合、各々の状況において、前述の三種類の視覚重畳関係（独立あるいは接触・片側遮蔽・相互重畳関係）が観測される。ここで改めて注目すべきは、相互重畳関係は何かがある場合（突入関係・一部包含関係）のみで観測される関係であることである。これは 2 章で図 2.4 に表した Op. 7, Op. 8 の成功状態を観測するのに、適している視覚的手がかりだということである。また、接触面に垂直に近い視点を選択し、かつ操作対象同士が近接・接触していることが分かれば、片側遮蔽関係の観測によって対象同士の上下関係が推定できる。

さらに、このように重畳関係を認識できたならば、通常は遮蔽により観測の難しい視線方向の対象同士の接触関係は、重畳関係そのものを使って修正し、一方で接触面に沿うような操作対象同士のズレは、対象の位置が分かるならば、直接使えば良いことになる。

このような考え方にに基づき、各操作で起きうる失敗の認識方法、そして失敗復帰方法をまとめて擬似コードにしたものを示す。Alg. 1 は Op. 1 から Op. 4 までの操作戦略、Alg. 2 は Op. 5 と Op. 6 のための操作戦略を示したものである。この操作戦略のコンセプトは次のようである：

1. 接触面の階層方向の成功・失敗は、シンボリックな階層関係として視覚重畳関係に基づき認識し、階層関係に応じて操作戦略の勾配方向を変えて対応する。
2. 接触面に平行なズレは直接観測した対象位置同士のズレを用いる。

ただし Θ , Θ_0 を方策と最初の方策とし、 δ , δ_0 をそれぞれ観測された対象同士のズ

レ、そしてズレの閾値とする．また $\varepsilon, \varepsilon_0, \varepsilon_{min}$ をそれぞれ戦略を変えるための方策の勾配，最初の勾配，そして最小の勾配値とする．また，勾配は学習係数 η_* によって制御するものとする．これは失敗復帰のための方策であるが，動作に必要なパラメタを決定していく側面からは，動作学習をしているとも言える．このように操作結果によって勾配方向を決定していく学習法としては方策勾配型強化学習 (PGRL) [43, 141] がある．本実験のような状況では一度でも試行が成功すれば良いことから，各試行ごとに方策をランダムに振らず，一回のみ試行していることが，PGRL とは異なる点である．これらはいずれも簡潔なアルゴリズムであるが，動作記述・認識のシンプル化に貢献している．本章次節の実験では，この動作アルゴリズムが実際の生活支援タスクにおいて有用であることを実証する．

Algorithm 1 Algorithm for Op. 1 - Op. 6

```

01:  $\Theta \leftarrow \Theta_0$ 
02: while true do
03:   executeTask( $\Theta$ )
04:   S,  $\delta \leftarrow$  checkState()
05: #if Op. 1, 3, 5
06:   if S = “One-way, down” and  $|\delta| < \delta_0$ 
07:     then return
08: #else
09:   if S = “Independent” and  $|\delta| > \delta_0$ 
10:     then return
11: #endif
12:    $\Theta \leftarrow \Theta - \eta \cdot \delta$ 
13:   endif
14: end while

```

Algorithm 2 Algorithm for Op. 7 and Op. 8

```

01:  $\Theta \leftarrow \Theta_0, \varepsilon \leftarrow \varepsilon_0$ 
02: while true do
03:   executeTask( $\Theta$ )
04:   S,  $\delta \leftarrow$  checkState()
05:   if S = “Mutual” and  $|\delta| < \delta_0$ 
06:     then return
07:   else if S = “One-way, down” then  $\Theta \leftarrow \Theta + \varepsilon - \eta_\delta \cdot \delta$ 
08:   else if S = “One-way, up” then  $\Theta \leftarrow \Theta - \varepsilon - \eta_\delta \cdot \delta$ 
09:   endif
10:    $\varepsilon \leftarrow \max(\varepsilon_{min}, \eta_\varepsilon \cdot \varepsilon)$ 
11: end while

```

4.6.2 自らが操る道具を追跡するための対象動作モデル

本節ではロボットが自ら道具を操作している状況で、道具と操作対象の関係を推定する場合を扱う。用いる視覚機能は本章で述べてきた手法と同様であるが、道具と操作対象をロボット自身が操作している場合は、ロボット自身の姿勢情報を用いることができるため、動作モデルは以下のようにできる。

$$\mathbf{x}_{t|t-1}^{(k)} = \mathbf{x}_{t-1}^{(k)} + d\mathbf{x}_t. \quad (4.13)$$

ただし $d\mathbf{x}_t$ は、ロボットのエンコーダ値から計算された、エンドエフェクタのカメラ座標系での速度である。

4.7 見えない対象の状態推定に基づく道具利用実験

実験には、3章で説明した広視野高視力ヒューマノイド HRP-2VZ を用いた。入力画像は、HRP-2VZ の高視力視覚（光学ズームカメラ）を用いたものである。

4.7.1 実験1：ファイルのフォルダ挿入実験

実験環境と実験条件

実験1の環境と条件を図4.13に示す。ロボットは図4.13のAに示すように、右手に樹脂製ファイル、左手に紙製のフォルダを持って、フォルダへファイルを挿入する。実験に使用したフォルダ二つとファイルを図4.13のBに示した。樹脂製ファイルは赤色、紙製フォルダは緑色で、同種二個を交代で実験に用いる。赤色のファイルの大きさは縦270[mm] × 横190[mm]、厚さは1[mm]未満である。また、緑色のフォルダの大きさは縦310[mm] × 横240[mm]、開口部のサイズは変化するが概ね0[mm]（全く隙間が見えない）から10[mm]の間である。ファイルとフォルダはともに、材質に起因する柔軟性を有する。

ロボットはまず、右手のクリアファイルを Δr [°] だけ下へ傾けた後、フォルダの方へ動かして挿入を試みる。しかし操作対象の柔軟性や開口部サイズの変化に伴って、操作には失敗が容易に起こり得る。この操作は2章で述べた操作 Op. 7 に分類

されるため，Alg. 2 を用いて失敗復帰をおこなった．失敗復帰の手法は以下のようである：

1. ロボットはまず右手を下に曲げ（図 4.13 .C.1），両手を近づけるようにしてファイルをフォルダへ挿入する試行をおこなう（図 4.13 .C.2）．
2. 試行終了時のファイルとフォルダの関係を，視覚重畳関係に基づき推定して，成功判定をおこない，
3. 挿入操作に失敗した場合には，視覚重畳関係に基づき Δr [°] を変更して操作をやり直す．

つまり，Alg. 2 で記述したパラメタ Θ を，右手を曲げる角度 Δr [°] としたものである．

ロボットの挿入操作後に，ファイルとフォルダが取り得る状態を図 4.14 に示す．状態は3つであり，それぞれ操作の目標となる“s (= Mutual)”状態，失敗であってファイルが上にずれてしまった“u (= One-way, up)”状態，そしてファイルが下にずれてしまった“d (= One-way, down)”状態である．

しかし，作業が成功するかどうかは紙製のフォルダの個体差により，あるいはフォルダの形状が変化するために，同じフォルダでも成功に必要なパラメタは異なる．ロボットがファイルをフォルダに入れるタスクを，右手首の回転角度 Δr [°] を変えながら行った場合に，どのような状態になるかを調査した予備実験結果を表 4.2 に示す．予備実験では Folder 1，Folder 2 のそれぞれに対し， dr を 1[°] 毎に変化させながら二回試行を行った．その結果，各フォルダの微妙な形状の違いによって，成功状態“s”となる角度が異なることが分かった．さらに，同じ角度で試行を行った場合でも操作後の状態が異なる場合があった（Folder 1, $dr = -9$ [°]）．

実験結果

全ての実験において，重畳関係を用いたファイルとフォルダの状態推定は正しく行われた．この間の視覚認識回数は 67 回であった．失敗復帰実験において，各種パラメタは $\Theta_0 = \delta r_0 = -6$ [°]， $\eta_\varepsilon = 0.5$ ，そして $\varepsilon_0 = 3$ [°] のように定めた．Folder 1 を用いた実験では，五回の実験全てで試行一回目にファイル挿入タスクが成功し，状

態“s”となった。一方，Folder 2 を用いた実験では五回の実験全てで，試行一回目には状態“d”となった。試行二回目ではいずれの実験でもタスクが成功し，状態“s”となった。Folder 2 を用いた実験の様子を図 4.15 に示す。フォルダへのファイル挿入タスクにおける平均試行回数は 1.5 回，成功率は 100[%] となった。

この操作の視覚認識においては，フォルダのインデックス部分の存在が，相互重畳関係を推定するために重要となっている。しかしインデックス部分が存在しなくても，より高解像度の画像を用いれば，フォルダの開口部のずれや厚みが完全にゼロでない限り，同様の推定が可能であると考えられる。また，認識時にマニピュレーションを利用し，裏側からも観測することで相互重畳関係を推定する手法も考えられる。

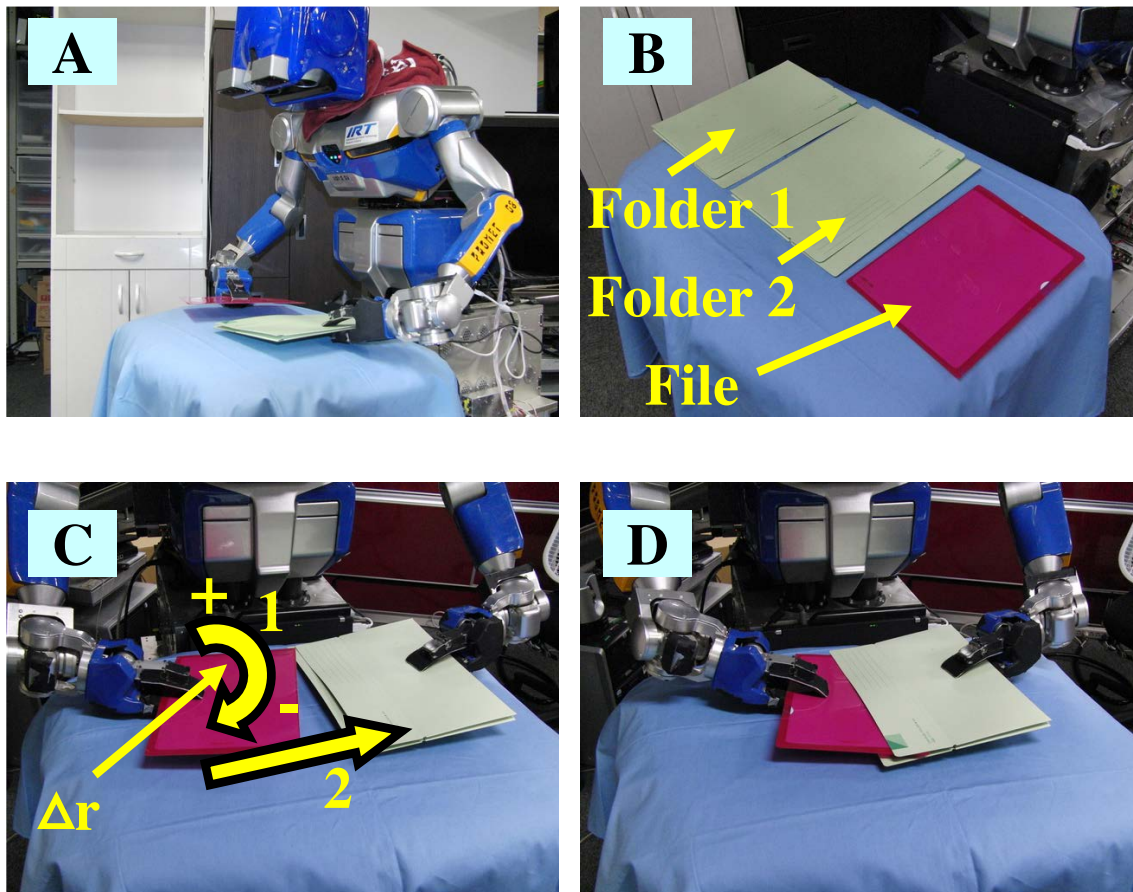
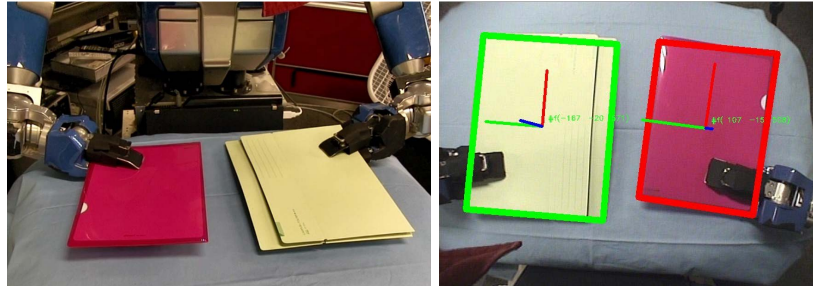


図 4.13: 実験 1 (ファイリング) の実験環境 . A) ロボットは右手でファイルを持ち , 左手でフォルダを持っている . B) 実験に使用したのは , 同種二個のフォルダと一個のファイルである . C-1) ロボットは右手を Δr だけ下へ傾けた後 , C-2) 両手を近づけてファイルをフォルダに入れる動作をおこなう . D) 成功時の状態 (著者論文 [106] より)

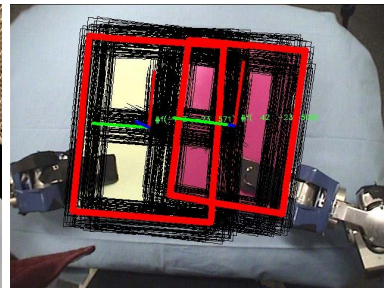
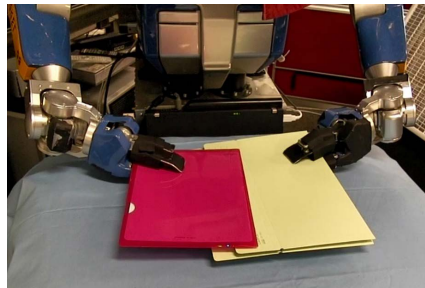
Initial state



State “s”

 $N_T = 11427$
 $N_o = 3733$


State “u”

 $N_T = 319$
 $N_o = 10695$


State “d”

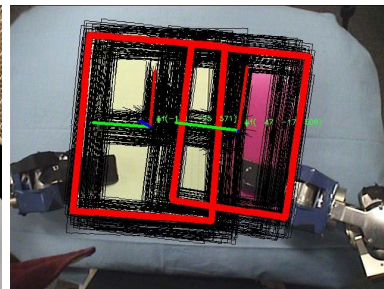
 $N_T = 14084$
 $N_o = 0$


図 4.14: フォルダへのファイル挿入動作試行後のフォルダとファイルの三状態 . State “s” (相互重畳) はファイルがフォルダに入っている状態である . また , state “u” (片側遮蔽 , 上) および state “d” (片側遮蔽 , 下) はフォルダにファイルが入らず , それぞれ上または下にずれ , 挿入に失敗している状態を表す . 下段の図は , ロボットの視野画像に , ファイルとフォルダの追跡結果を赤線でオーバーレイしたものである (著者論文 [106] より)

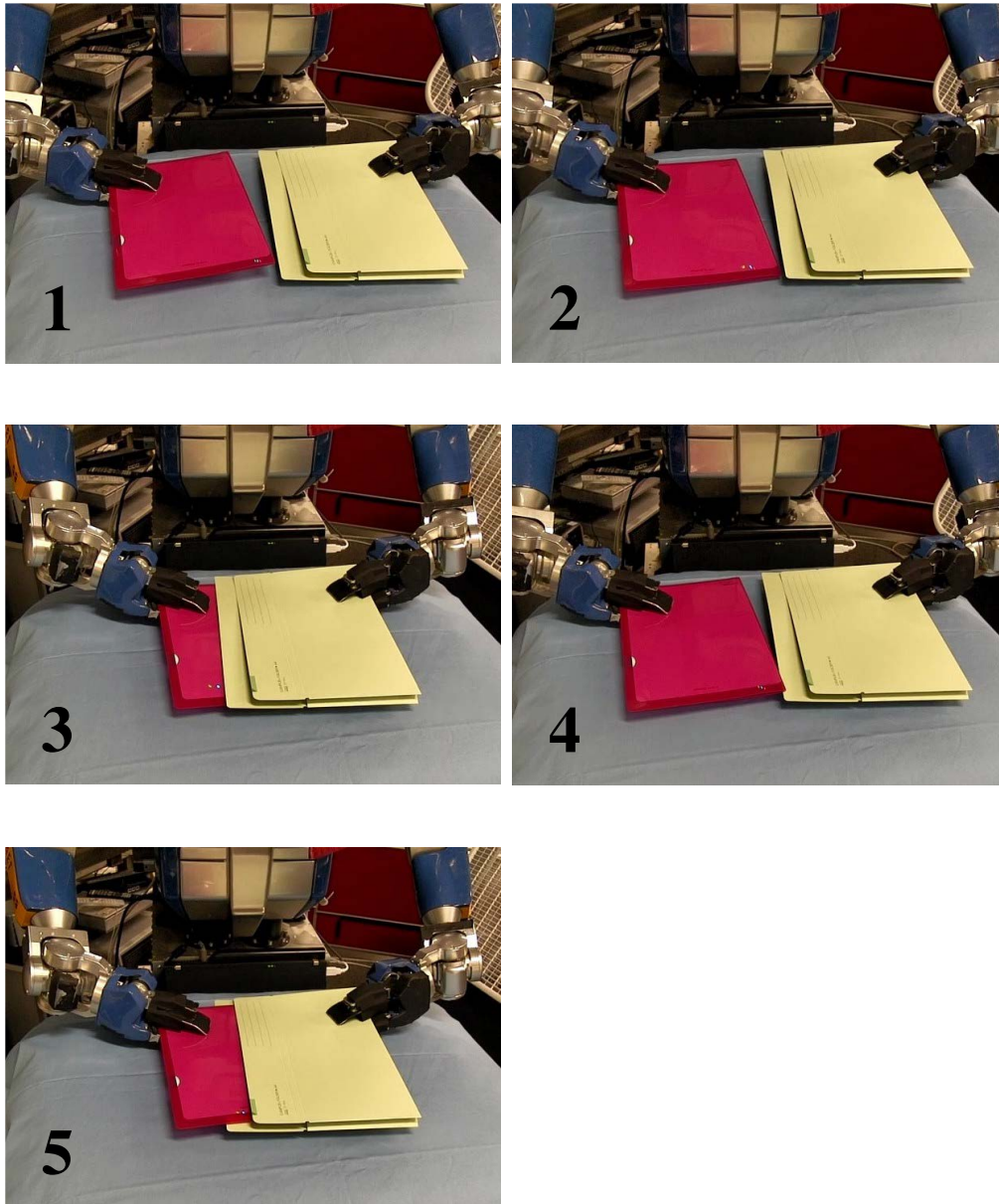


図 4.15: 実験1 (ファイリング) の結果. 挿入動作に一度失敗した後 (3), パラメタを変更して失敗復帰動作がおこなわれ (4), ファイリング動作に成功した (5) (著者論文 [106] の図を改変)

4.7.2 実験2：目玉焼きのよそいつけ作業

実験環境・実験条件

2章図2.3のBに示したのが、目玉焼きよそいつけ作業の実験環境である。ロボットは右手でフライ返しを持ち、フライパンの中にある目玉焼きをすくって持ち上げ、右側の皿によそいつける。フライ返しの持ち方、目玉焼きと皿の位置は定まっているものとする。

しかしこのタスクでは、フライ返しを目玉焼きの下に差し込む動作がうまく行ったり行かなかったりする。つまり、フライ返しの上にきちんと目玉焼きが載る場合と載らない場合がある。図4.17に、(A)成功した場合と(B)失敗した場合の実験の様子を示す。(A)では成功しているが、(B)ではフライ返しの上にうまく目玉焼きが載らず、フライ返しを上を動かした際にフライパンの外へ落としてしまった。

そこで、目玉焼きをすくう動作に失敗復帰の導入を検討する。フライ返しを差し込む動作後の状況(A)、(B)それぞれの、フライ返しを差し込む直前と直後のロボットの視野画像を図4.16の1, 2に示す。視野画像でも、失敗した場合(B-2)ではフライ返しの上に目玉焼きがきちんと載っておらず、成功した場合(A-2)では載っていることが分かる。このような操作は2章で述べたOp. 5に分類されることから、失敗復帰にはAlg. 1で述べた手法を用いることが可能である。フライ返しと目玉焼きを、それぞれ2章で述べたObject A, Bとし、以下のように操作をおこなうものとした。

1. フライ返しを目玉焼きの下に差し込む動作をおこなう。
2. 動作終了後、フライ返しが目玉焼きに対して“One-way, down”の状態であることを確認する。重要な接触面に水平な方向の操作誤差 (misalignments, $|\delta| < \delta_0$) には、目玉焼きとフライ返しの重なっている領域のサイズを利用する。重なり領域が1000ピクセル以上であることを、すくう操作の成功状態であると定め、失敗判定をおこなう。
3. 失敗復帰動作では、フライ返しを元の位置に戻した後、フライパンを振って目玉焼きがフライパンの中心部に戻るようにし、もう一度すくう操作をし直すという操作をおこなう。

4. 五回連続で失敗復帰に失敗した場合には、動作を停止する。

変形する対象領域の推定について

この実験における目玉焼きのように、個体毎に形状が異なったり、柔軟性により形状が時々刻々変化するような対象においては、先に述べたような、テンプレートを透視変換するだけでは領域を適切に追跡することができない。そこで本実験では、前述の色ヒストグラムによるグラフカット法を用いて卵の領域を推定するものとした。必ず入りうる領域を種とするように TRIMAP を作り、グラフカットにより領域境界を決定する。

道具の追跡の際には、道具を動かした後の時刻 $T + 1$ における道具側の遮蔽予測領域を、道具を動かす前の時刻 T における目玉焼き領域と、幾何モデルから計算した、時刻 $T + 1$ における道具領域の非重畳領域として計算するものとした。

実験結果

はじめに、ロボットに前述の視覚機能・失敗復帰機構を持たせずに、目玉焼きのよそいつけ動作をさせる基礎実験をおこなった。この基礎実験では、全 14 試行のうち目玉焼きのよそいつけに成功したのは 4 試行、成功率は 29[%] であった。

次に、視覚状態推定機構の単独評価実験をおこなった。一部の結果を図 4.16 に示す。まず、目玉焼きをフライ返しに載せることに成功した際の画像列を入力として 10 回、視覚状態認識をさせたところ、全 10 回で “One-way, down” の成功状態が推定され、認識成功率は 100[%] であった。次に、目玉焼きをフライ返しに載せられなかった場合の画像列を入力としたところ、全 34 回の試行のうち、32 回の試行において視覚重畳関係は “Independent” の失敗状態だと推定された。成功率は 94[%] である。

さらに、この視覚機能と失敗復帰機構をロボットへ実装して、目玉焼きをよそいつける実験をおこなった。結果、全 10 回の試行のうち、最終的によそいつけに成功したのは 6 回であった。図 4.18 に、成功した試行のうち一つの実験結果を示す。よそいつけに失敗した 4 回の試行のうち、ロボットが目玉焼きをフライパンの外へ落としたのは、1 回であった。残りの 3 回の試行は、ロボットは 5 回連続で失敗復帰動

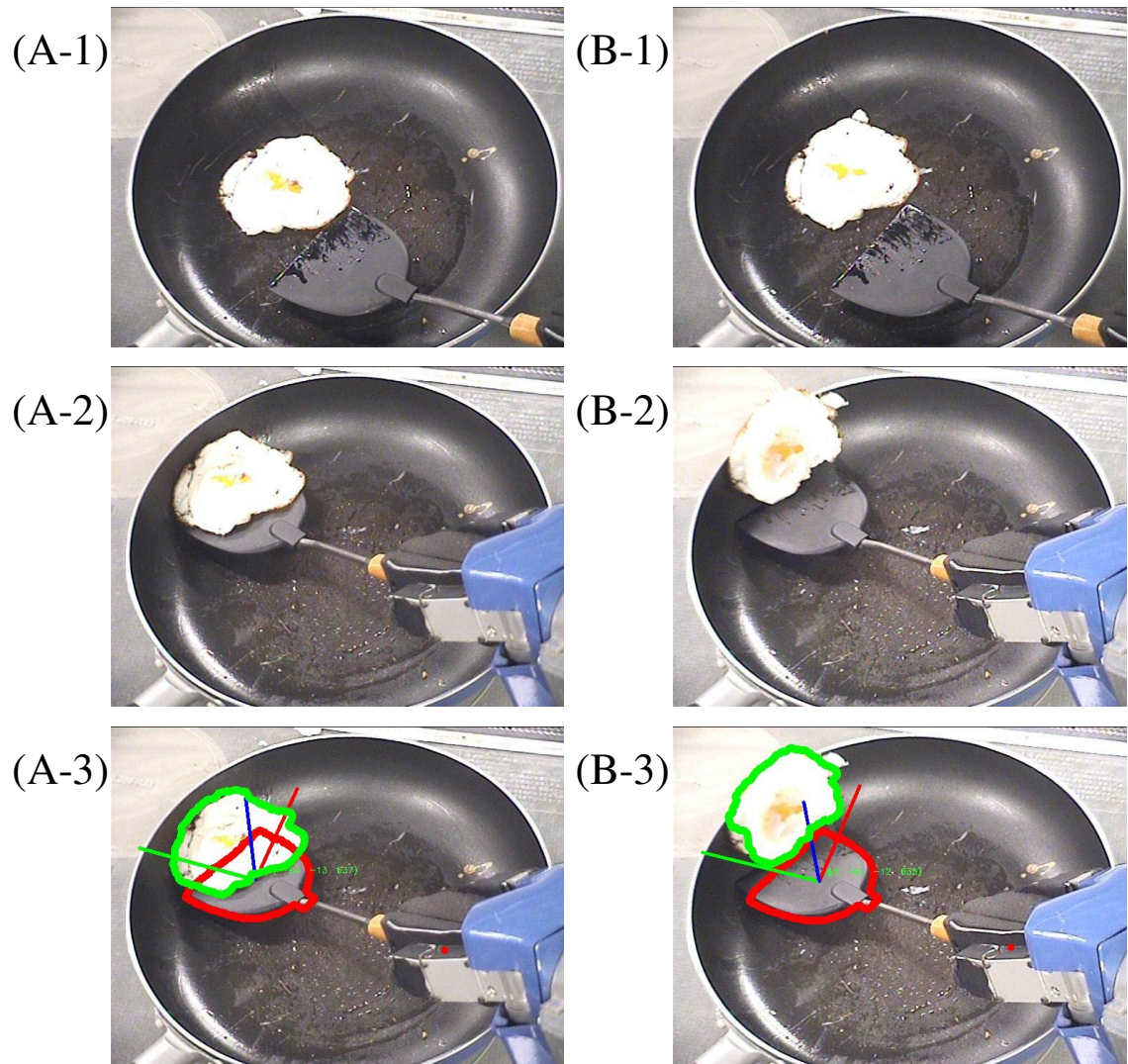


図 4.16: 目玉焼きの操作に (A) 成功している場合と (B) 失敗している場合の視覚画像と追跡結果 (著者論文 [54] の図を改変). 緑で囲われた領域が目玉焼き領域の推定結果, 赤で囲われた領域が道具領域の推定結果, 赤・緑・青の線分は道具座標の各々x軸・y軸・z軸を表す

作に失敗したものである。この原因は、人が作成した失敗復帰動作がうまく働かなかったことである。フライ返しを目玉焼きの下へ差し込む前に、目玉焼きが元の位置へ戻らなかったことが主な要因であり、この解決のためには、失敗復帰動作を変更していくことだけでなく、あるいは失敗復帰動作をロボット自身に変更していく機構を構築していくことが重要だと考えられる。最終的に、ロボットが目玉焼きを落とさずに操作する、という失敗しない確率は90[%]であり、提案した状態認識と失敗復帰手法の有用性を示している。

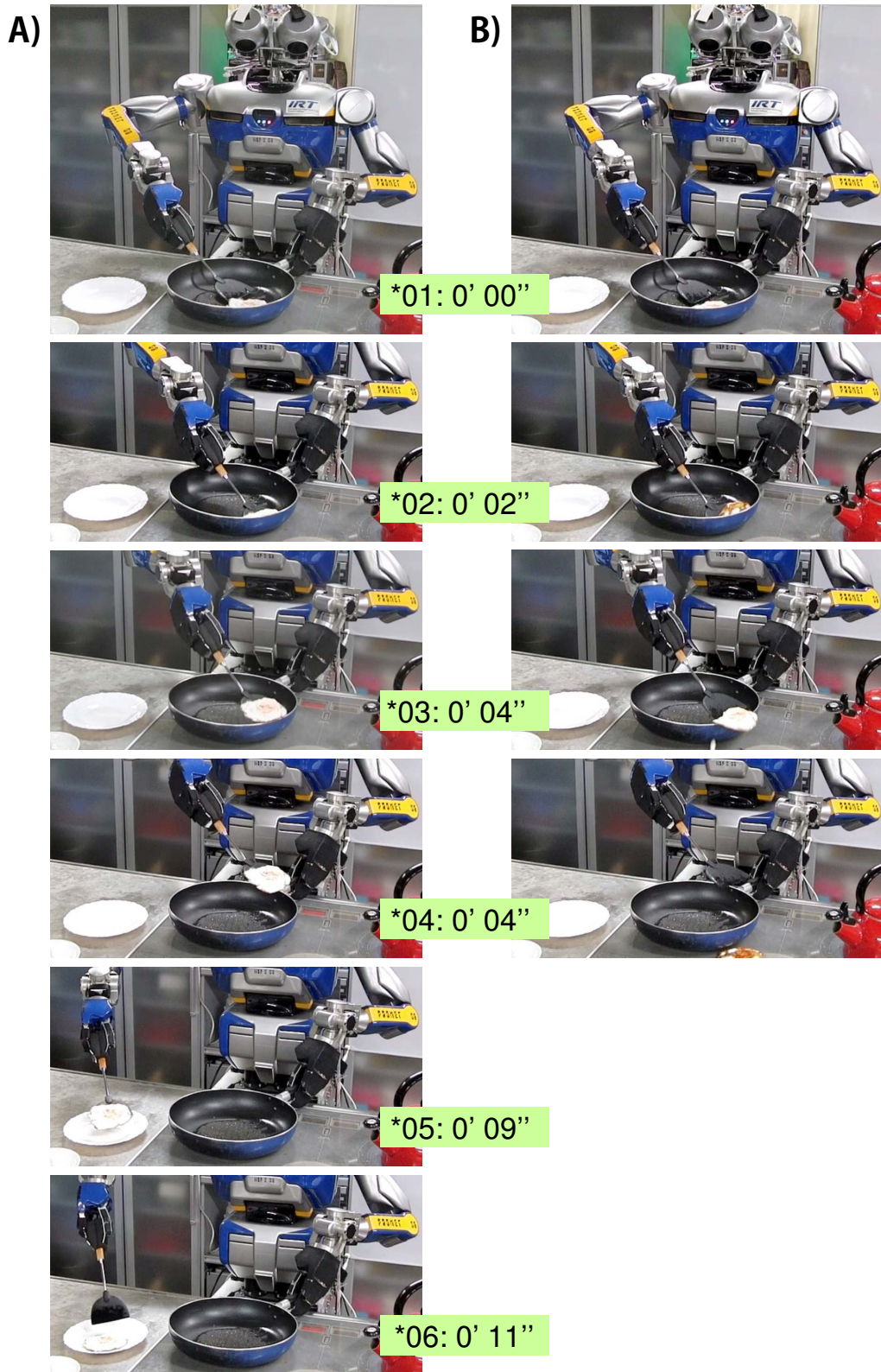


図 4.17: 目玉焼きよそいつけ動作における A) 成功と B) 失敗 (著者論文 [54] の図を改変)

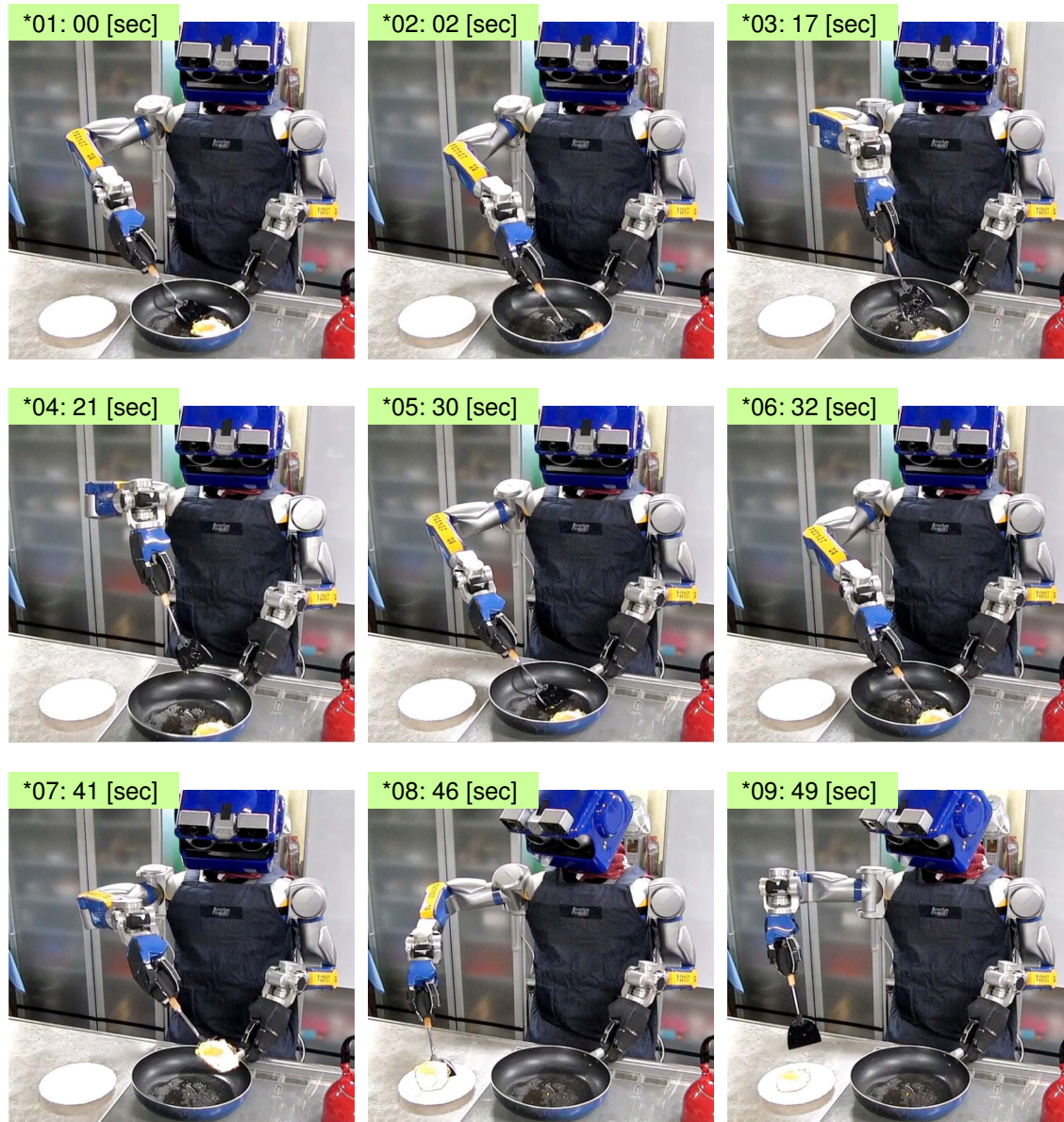


図 4.18: HRP-2VZ による，“LOAD” 判定を用いた失敗復帰に基づく，料理よそいつけ実験結果．(1) トングを目玉焼きの下に差し込もうとするが，(2) 失敗する．視覚処理結果が失敗と出たので，(3-4) 失敗復帰を行う．トングを一度戻してフライパンを振り，目玉焼きを元の位置に戻す．(5-6) もう一度トングを目玉焼きの下に差し込む．今度は視覚処理結果が成功と出たので，(7) 目玉焼きを持ち上げて，(8-9) よそいつけに成功した（著者論文 [54] の図を改変）

4.8 おわりに

本章では、生活支援ロボットが人を観察する際、そしてロボット自身が道具を利用する際における、道具と操作対象の間にみられる作用、特に上下・突入関係の変化を伴う作用について視覚推定するための手法について述べた。この手法では、遮蔽下の状況を周囲の空間情報（空間遮蔽近傍情報）を利用して推定する。部分的に、しかし大規模に遮蔽が発生するような、テクスチャの少ない対象について、非遮蔽予測領域のみを尤度計算に用いて追跡をすることは、遮蔽部分を含む対象位置の同定に有用であった。対象同士に上下・突入・部分的な包含関係が発生している場面でも、遮蔽下の道具領域を適切に推定することが可能であることから、重なり領域を特定し、対象領域同士の重畳関係を推定することが可能であった。これは2章で問題提起した「日常的な道具と操作対象の間の遮蔽関係の観測の可能性」を肯定するものである。

また、視覚重畳関係のうち片側遮蔽関係を用い、運動の随伴性情報を併用して接触関係を確認することで対象同士の支え動かす関係を、相互重畳関係を用いることで、平面状の道具が凸面に切り込む、あるいは凹状の対象に他の対象が入り込むような作用の推定が可能であることを実証的に示した。さらに、ロボット自らが動かしている薄い対象同士の上下関係推定には、片側遮蔽関係をそのまま用いることが有用であった。これは2章で問題提起した「視覚重畳関係を用いた物体同士の関係推定の可能性」を肯定するものである。例えば観察対象の表面法線方向を利用することで、「混ぜる」と「切る」の分割が可能であることなどが考えられ、本手法を基軸として、対象同士の関係や関係の変化を表す述語を増やしていくことは可能だと考えられる。このように得られた結果は、観察した道具の機能を学習すること、道具と操作対象の組み合わせを学習することに貢献するものである。

さらに、ロボットの動作の失敗復帰において視覚重畳関係計算するシステムを用いることは、ロボットが従来観測しにくかった、薄い対象同士の上下関係や重なりのおおきい対象同士の関係推定へ有用であることを実証的に示した。この実験においては、ロボットは大部分が遮蔽されている対象を追跡できるだけでなく、視覚重畳関係を用いて、従来は推定しにくかった方向へも軌道を修正することが可能である。対象同士が近接・接触して作用し合うタスクについても、本システムが軌道レベル

での学習に有用であることを指摘し，2章で述べた「視覚重畳関係を用いた物体同士の関係推定を用いた動作学習の可能性」を肯定するものであると言える．本視覚追跡手法によると，近接・接触し合う道具と操作対象が重度の遮蔽下にあっても，それらの軌道を得られるものであったことから，この結果は軌道レベルでの道具操作学習一般へ活用可能だと言える．

さらに「生活支援タスクの学習・実行に際しての有用性」に関しても，ロボットが目玉焼きをよそいつける，ファイル挿入作業をおこなうという，それ自身に問題点が存在し，従来法では難しかったタスクを実現可能であったことから，本章で述べた視覚機能と失敗復帰法は生活支援タスクへ有用であることを実証的に示したと考えている．

ただし本手法は，遮蔽下にある対象の状態を，各々の時刻における非遮蔽下の領域を用いて推定するものであった．何かを何かに完全に入れる等の，対象領域が完全に遮蔽下となる状況においては，その時刻以外の情報をも利用する必要があると考えられ，次章でその手法について検証する．

第5章

遮蔽時刻近傍情報を用いた見えない作用推定に基づく学習法

5.1 はじめに

前章では、隠れている道具や操作対象の状態を推定するために、遮蔽されている空間近傍の特徴を用いることで、遮蔽されている状況を推定するための手法について論じた。

この手法は観察対象の少なくとも一部が見えていることを前提としている。一方で、上下・包含関係の変化によって道具や操作対象全体が完全に遮蔽されてしまうような状況や、フレームアウト等で認識の失敗がたびたび含まれるような状況も起こり得る。例えば物体が布袋に入れられた後に運ばれるようなシーンや、トレーに大きな物体を載せて運ぶようなシーンであるが、空間遮蔽近傍情報が存在しないため、前章の手法をそのまま利用することはできない。そこでこのような場合には、観察対象が遮蔽される時刻の近傍の見え方のみを使って、遮蔽下の観察対象の状況を推定する枠組みが必要となる。

本章では、観察対象の認識に成功した時刻のみの認識結果、つまり遮蔽時刻近傍情報を用い、観察対象の認識に失敗する状況の分類に基づいて各時刻の観察対象の遮蔽下・非遮蔽下での状態推定をおこない、上下・包含関係の変化を含む作用がいつおこなわれているかを推定することが可能か、検証していく。具体的な状況として、片付けや荷物の仕分け等を考えると、タスク目的としては、物体をどこへ移動すべきか、物体を何に入れるべきか、何に載せるべきか、という情報が重要だと考えられる。さらに、観察対象の道具的利用法としては、何かを上に乗せて良いか、何かを上に乗せた状態で運んで良いか、何かを中に入れて良いか、何かを中に入れた状態で運んで良いか、という利用可能情報が考えられる。物体を載せたり入れたりして運ぶこと、一時的に物体を置いておく場所として利用することは、効率的な物体移動タスクの実現へ貢献するものの、物体の破損や汚れにつながる可能性も有する。しかし、人間が家庭内で普段利用している際の道具的利用法が推定できれば、ロボットもその範囲内で物体を道具として利用し、効率的なタスク再現が可能となるであろう。

そこで本章では、観察対象が正しく認識できている時刻のみの認識結果を用い、上記のタスク目的と道具的利用法を推定するシステム“V-FES”(Visual Hierarchy-based Function Estimator)を提案し、評価するものとする。



図 5.1: 生活支援ロボット HSR による片付け観察学習環境 (著者論文 [104] の図より)

従って本章で検証するのは，2章で述べた以下の事項である．

1. 日常的な道具と操作対象の間の遮蔽関係の観測の可能性
2. 視覚重畳関係を用いた物体同士の関係推定の可能性
3. 視覚重畳関係を用いた物体同士の関係推定を用いたタスク学習の可能性

なお本章の記述は，著者論文 [100]，[104]，および出願済特許 [127] の内容を含む．

5.2 視覚重畳関係を利用する V-FES の構成

前節で述べたように、本章で取り上げるタスクは、箱に物をしまう・トレーで物を運ぶ等の、道具に物を重ねる・入れる、そしてその状態で物を運ぶようなタスク(図5.1)である。このようなタスクでは、例えば物に物を重ねる場合や入れる場合に、観察対象がカメラに映らない状況が発生する。また、観察中に対象がフレームアウトする状況も容易に起こり得ることから、どのような追跡手法を用いても、観察対象を正しく認識する、つまり位置を正しく推定することは難しいと言える。

2章で指摘した通り、一般に観察対象が認識できなくなる理由としては、【NON-VISIBLE-1】センサの観測領域に含まれない場合(観測不可能)【NON-VISIBLE-2】センサの観測領域に含まれているが、他の物体に遮蔽されている場合(観測不可能)【VISIBLE+UNRECOGNIZED】観測可能だが、物体を特定するのに十分な視覚特徴がない場合(小さすぎる、見た目の区別がつかない等)が考えられる。また、認識された場合でも、認識結果が誤っている場合として【VISIBLE+WRONGLY RECOGNIZED1】本当は観測不可能であるのに、誤認識してしまう場合、【VISIBLE+WRONGLY RECOGNIZED2】観測可能だが、誤った場所に認識してしまう場合、が考えられる。観察学習の際には、対象の道具的な利用によって【NON-VISIBLE-2】が起こる。また観察者のセンサの動きが追いつかない場合に、フレームアウトによって【NON-VISIBLE-1】が、対象との距離が遠くなることで【VISIBLE+UNRECOGNIZED】が容易に起こり得る。

改めて遮蔽下にある対象の認識法について考察すると、対象のテンプレートを変形しながら追跡することで、部分的な遮蔽に対応する手法[37, 53]、ある程度の遮蔽に対応した特徴量を用いた探索手法[46]等が研究されている。しかしこれらの手法では、本章のタスクで発生する【NON-VISIBLE-1】や【NON-VISIBLE-2】の完全な遮蔽には対応できない。遮蔽下では周辺の視覚特徴を記憶して追跡する手法[30, 119]も研究されており、完全な遮蔽下でも追跡を可能としているが、完全な遮蔽下での大きな移動やフレームアウトが起こると、追跡は不可能だと言える。

本章では【NON-VISIBLE-1】から【VISIBLE+UNRECOGNIZED】へ対応するため、大幅な遮蔽の後でも対象の位置を特定できるよう、SIFT[46]やLINE-MOD[31]等の探索手法の認識結果を入力として用いる枠組みとする。さらに【VISIBLE+WRONGLY

RECOGNIZED1】と【VISIBLE+WRONGLY RECOGNIZED2】に対応するため、探索手法の検出の評価を厳しくし、偽陽性のない認識結果を期待するものとする。これにより増加する【VISIBLE+UNRECOGNIZED】のミスについては、認識されない時区間を埋めるための知識を用いて対応する。そして探索成功時刻のみの認識結果を用い、各時刻の対象の遮蔽下・非遮蔽下での状態推定をおこない、いつ何の作用がおこなわれているか、タスク目的、そして対象の道具的機能を推定するものとした。

本研究で提案する V-FES を用いる観察学習システムの構成を図 5.3 に示す。V-FES の入力は、観察対象の探索に成功した時刻のみの認識結果（位置、図 5.3、右上）である。V-FES では 1) 対象の道具的利用法、2) タスク目的を推定し、その結果はタスクプランナへ送られる。タスクプランナでは STRIPS[25] タイプのオペレータを用いて、文献 [44] や [57] で検証されているシンボリックなタスク再現計画をおこなうことで、ロボットが物体に許される操作のみを用い、ロボット自身の体に合わせてタスクを再現するための動作列を決定するものとする。

V-FES で推定し、タスクプランニングで利用する対象の道具的利用法は、前節で述べたように

1. 他の対象を載せることのできる “MAT” 機能、
2. 載せた状態で動かすことのできる “TRAY” 機能、
3. 他の対象を入れることのできる “STORAGE” 機能、
4. 入れた状態で動かすことのできる “CONTAINER” 機能

の四種類とする。タスク目的としては、

1. 対象をどこへ移動するか、
2. 対象同士の関係をどのようにするか（上下・包含関係の変化）

を計算する。これにより、片付け等に代表される道具利用タスクの推定・記憶をおこなう。

各時刻で観察対象に何が起きているかを、V-FES で推定するためのコンセプトを図 5.2 に示す。まず、観察対象が動かされたことが分かり、移動後の場所が分か

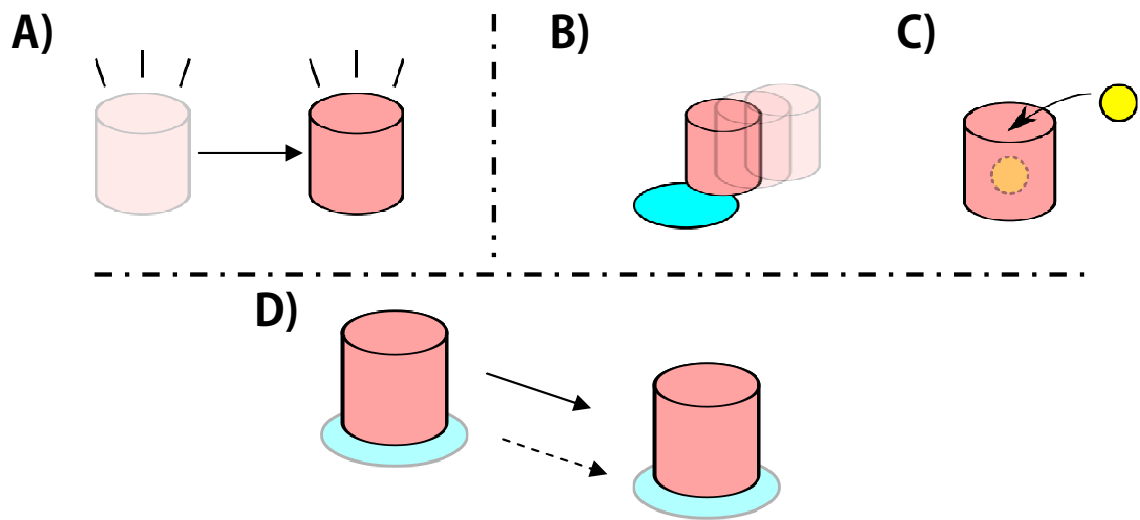


図 5.2: 上下・包含関係と移動を含む作用推定法の概要

れば、観察対象の移動の目的地が推定できると考えられる（図中、A）。次に、一方の観察対象に他方の観察対象が近づき、前者が見えなくなるような片側遮蔽関係が観測された場合には、前者の上に後者が載ったと推定する（図中、B）。また、一方の観察対象に他方の観察対象が近づき、後者が見えなくなるような片側遮蔽関係が観測された場合には、前者の中に後者が入ったと推定する（図中、C）。さらに、一方の観察対象が見えていなかったとしても、見えている観察対象が動き、見えていない観察対象も随伴して動いていることが分かれば、見えていない観察対象がどこへ動くべきかが推定できる（図中、D）。この推定のもと、中へ入れられた観察対象の移動と移動後の場所、あるいは上へ物を載せられた観察対象の移動と移動後の場所を計算する。

このような作用推定をおこなうためには、二つの観察対象の各時刻における視覚重畳関係と、運動特徴というメタな情報を推定する必要がある。しかし、観察対象が完全に遮蔽される場合、あるいはフレームアウトするような場合には、前章のような非遮蔽領域を用いて重畳領域を推定する手法が適用できない。そこでV-FESでは、各メタ情報の推定に際して、直接観測・計算できないメタ情報については、前後の時刻のメタ情報を利用して補間するものとした。ただし、視覚重畳関係と運動

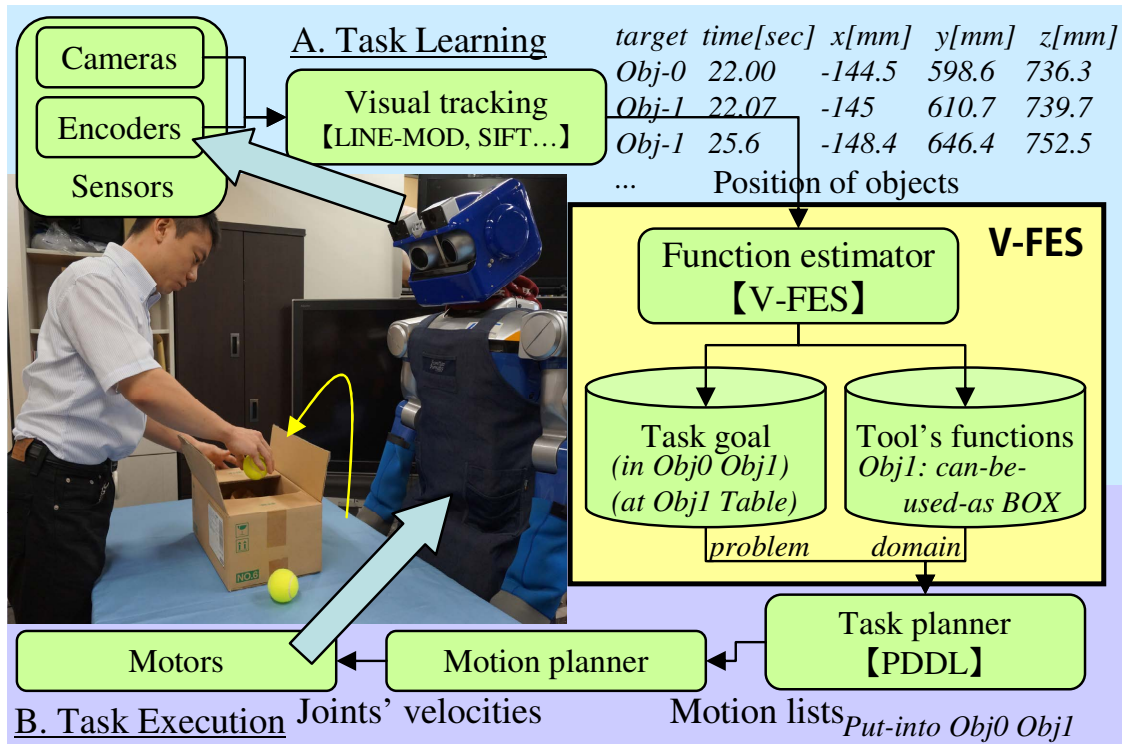


図 5.3: V-FES を用いたタスク学習の流れ (著者論文 [100] より)

特徴のみでは情報が足りないため，物体の視認性を含む五つのメタ情報（視覚の手がかり）を利用する．

このような V-FES の処理の流れを，図 5.10 に示す．五つのメタ情報の時系列データを，五つの知識 (KNO-1 から KNO-5) に基づいたアルゴリズムを用いて，順に補間・修正していく．最終的な時系列メタ情報から，各時刻でおこなわれている作用を推定し，前述の道具的利用法とタスク目的を算出するものとする．図 5.4 は，空間遮蔽近傍情報を用いた作用推定と，遮蔽時刻近傍情報を用いた作用推定を比較し，各知識の適用方向を示したものである．次節以降では，メタ情報と，メタ情報時系列データを修正していくためのアルゴリズムについて，順に説明していく．

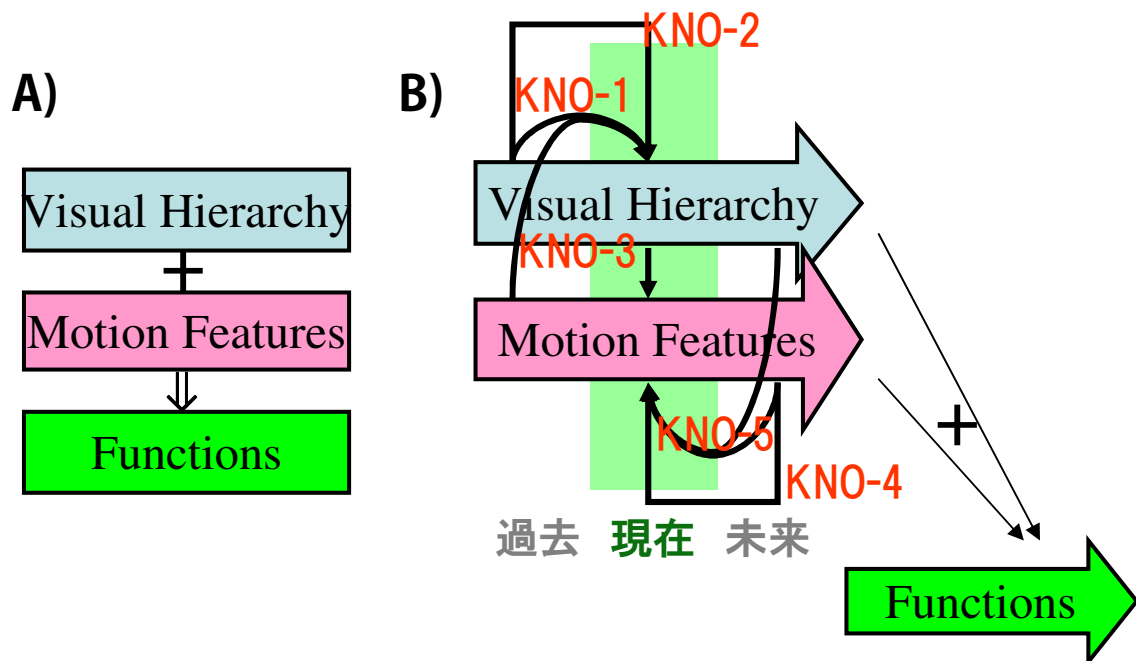


図 5.4: A) 空間遮蔽近傍情報を用いた作用推定と, B) 遮蔽時刻近傍情報を用いた作用推定における, メタ情報の流れ. B) には, KNO-1 から KNO-5 までの知識の適用方向を示した (著者論文 [100] の図を改変)

5.3 遮蔽時刻前後の特徴を用いた作用推定

5.3.1 作用推定のためのメタ情報

本研究で, 二つの対象間の作用の推定をおこなうための手がかりとするメタ情報は, 以下の通りとする.

- 対象の視認性

1. 認識の可否 Visibility (s_V^k)

各対象ごとに存在する状態 (小文字の s で表す) で, 対象 k が観測可能であるだけでなく, 認識器により検出がなされており, 位置が特定できている場合を “Recognized” (“Rec.”), それ以外を “Unrecognized” (“Unrec”) とする.

- 対象の運動特徴

2. 各対象の動きの有無 Movements (s_M^k)

各対象ごとに存在する状態で、対象が動いているときは“Yes”，動きが少ない場合は“No”とする．動きが分からない場合は“Unknown”とする．

3. 対象同士の距離 Distance (S_D)

二つの対象で一つ存在する状態で（大文字の S で表す），対象同士の距離が近い場合は“Near”，遠い場合は“Far”とする．距離が分からない場合は“Unknown”とする．

4. 対象同士の運動随伴性 Co-movements (S_C)

二つの対象で一つ存在する状態で、対象がもう一方の対象を動かしている作用の推定に用いられるメタ情報である．対象が動いている際に、対象同士の運動ベクトルが類似している場合を“Yes”，そうでない場合を“No”とする．分からない場合は“Unknown”とする．

- 対象同士の視覚重畳関係

5. 対象同士の視覚重畳関係 Visual Hierarchy (S_H)

二つの対象で一つ存在する状態で、対象同士の重なり方を表すメタ情報であるが、2章の考察に基づき、タスクに重要な接触・近接面から十分離れた場所から観測し、かつ近接関係を確認することで、物体を重ねる・入れるという真の物体の重なり方と近似できると考える．重なり関係が存在しない場合を“Independent”（“Indp.”），一方が全面にある場合（片側遮蔽）を“One-way”，特に対象 k が前面に存在する場合を“One-way (OBJ- k)”とする．

以後、これらのメタ情報の推移は、図 5.5 に示すような凡例色を用いて表現するものとする．

5.3.2 遮蔽が発生するシミュレーション実験の入力データ

本章で提案する V-FES の説明のため、そしてその基本的な能力の確認のため、まず片付けシーンにおいて意図的に抜けがある物体追跡シミュレーションデータを作成した．本小節では、このデータ列について説明する．

実験に使用した観察データセットは、図 5.6 に示す SIM-OBS1 と、図 5.7 に示す SIM-OBS2 の二セットである．SIM-OBS1 では、コップと皿を観察対象とする．コップは皿に重ねられ、両観察対象ともにキッチンへ運ばれる．SIM-OBS2 ではボール

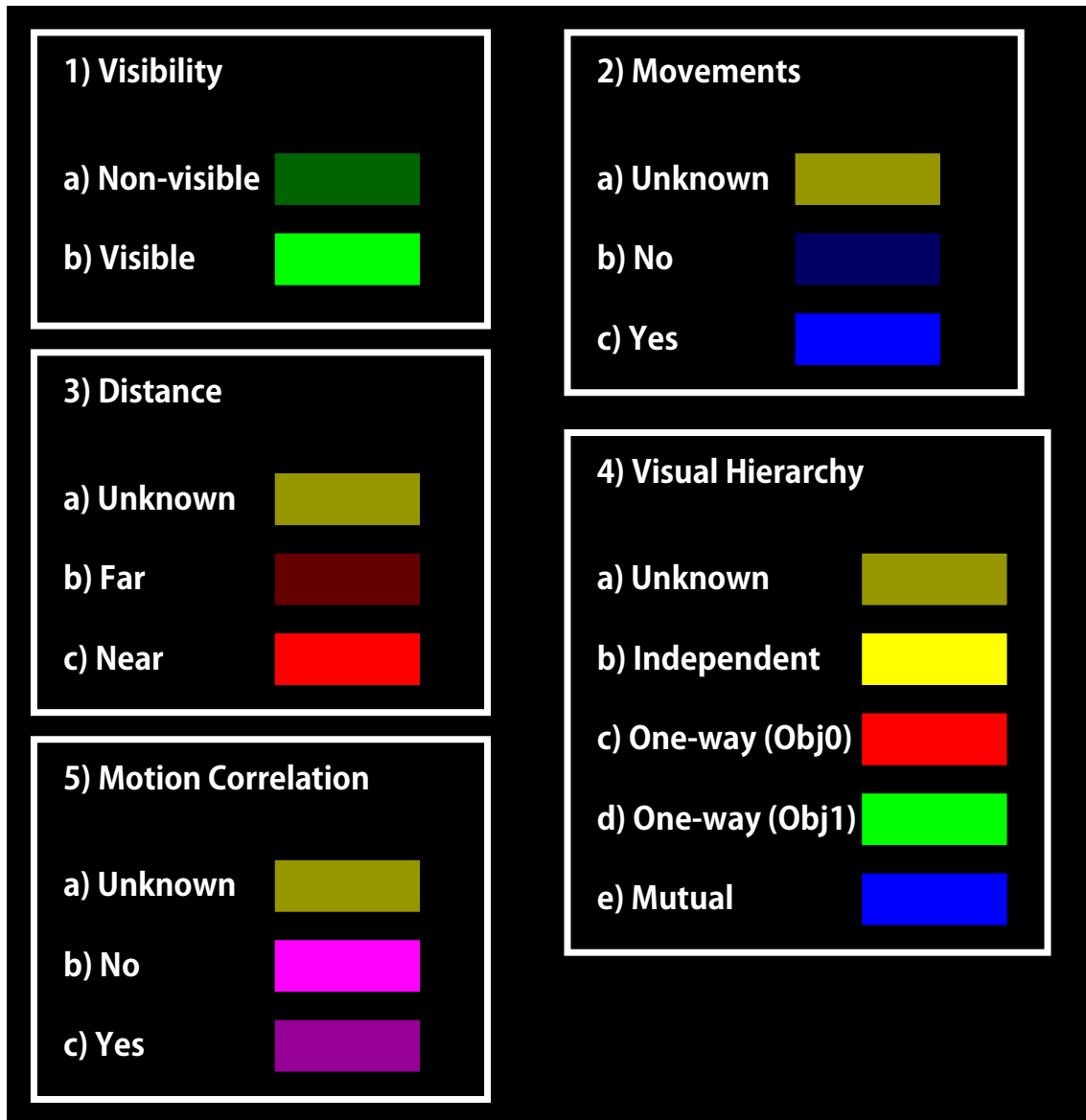


図 5.5: V-FES で用いるメタ情報の描画のための凡例 (著者論文 [100] の図を改変)

と箱を観察対象とする．ボールが箱に入れられ，箱が棚へと収納されるが，箱を運ぶ際に両観察対象ともに棚に隠れ，全く画像内に写らない時区間が存在する．

一般に画像中の対象探索では，観測可能な対象の特徴が多いほど探索結果は良好となる．対象の特徴が均一であれば，対象の観測可能な領域のサイズに依存する．そこで，時刻 t に対象 k が認識される確率 $P^k(t)$ を，対象 k の領域が隠れのない場合に画像中で取り得る面積 $A_{all}^k(t)$ と，実際に観測可能な領域の面積 $A_{rec}^k(t)$ を用いて，次式のように定めた．

$$P^k(t) = \begin{cases} 0, & \text{if } A_{rec}^k \leq A_{min} \text{ or } \frac{A_{rec}^k(t)}{A_{all}^k(t)} \leq r_{min} \\ \frac{P_{max}}{1-r_{min}} \left(\frac{A_{rec}^k(t)}{A_{all}^k(t)} - r_{min} \right), & \text{otherwise} \end{cases} \quad (5.1)$$

ただし， A_{min} は認識に最低限必要となる観測可能領域のサイズ， P_{max} は対象領域全てが観測可能だった際の認識成功率， r_{min} は認識に最低限必要となる対象の観測可能領域の割合を表す定数で $0 < r_{min} < 1$ とする．

$P_{max} = 0.7$ ， $r_{min} = 0.5$ ， $A_{min} = 400[\text{pixel}]$ として，対象の探索に成功するフレームをシミュレートした結果を，図 5.8 に示す．SIM-OBS1 ではコップを重ねたときの皿，SIM-OBS2 では箱に入れられたときのボールと，棚に隠れたときの箱が認識されていないことが分かる．

V-FES の目的は，図 5.8 に示されるような，観察中に観察対象が検出されない区間が顕著に存在する場合にも，検出失敗下での観察対象の状態を推定し，あるいは検出されないこと自体を利用して，観察対象間に表れる作用を推定するものである．V-FES の入力は，観察対象が見えている時刻のみの認識結果 (objectID, t , x , y , z) であり，一例を図 5.9 に示す．

5.3.3 追跡結果からの運動情報推定

まず，認識の成否 s_V^k を決定する．一定時間 t_C 毎に時区間 ($\mathbb{T}_{All} \in T_0, T_1, \dots, T_{fin}$) を設定し，各時区間に対象 k の認識結果が存在すれば，その区間で $s_V^k(T_i)$ は “Rec.”，そうでなければ “Unrec.” とする．対象 k ($k = 0, 1$) が時区間 T_i で “Rec.” であった場合，認識結果の三次元位置 $\mathbf{x}^k(t)[\text{mm}]$ ($t_C \cdot i \leq t < t_C \cdot (i+1)$) を平均してその時区間での位置 $\bar{\mathbf{x}}^k(T_i)$ とする．対象が一度も検出・認識されなかった時区間では，平

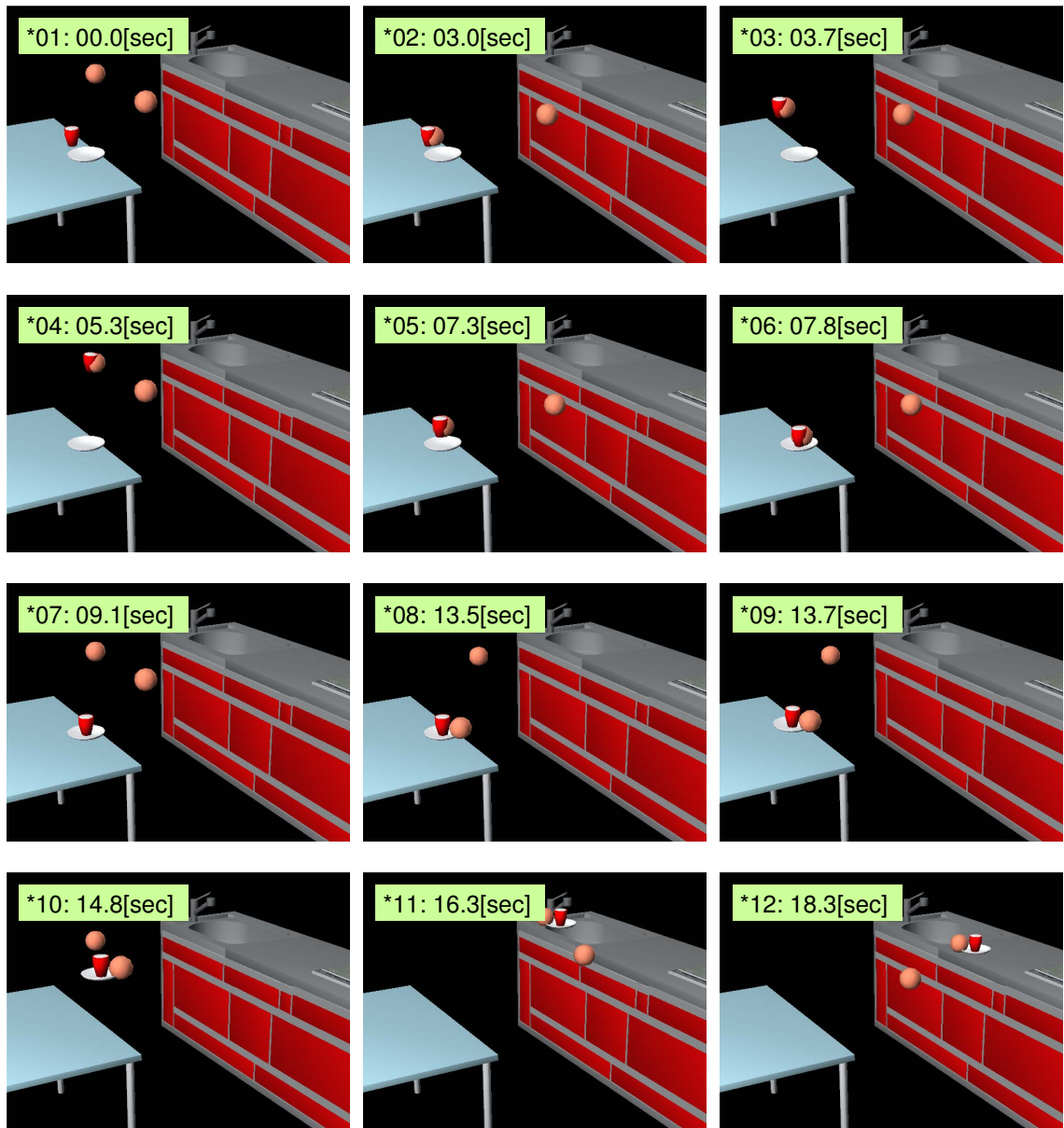


図 5.6: 食器を片付ける観察画像列 (著者論文 [100] の図を改変)

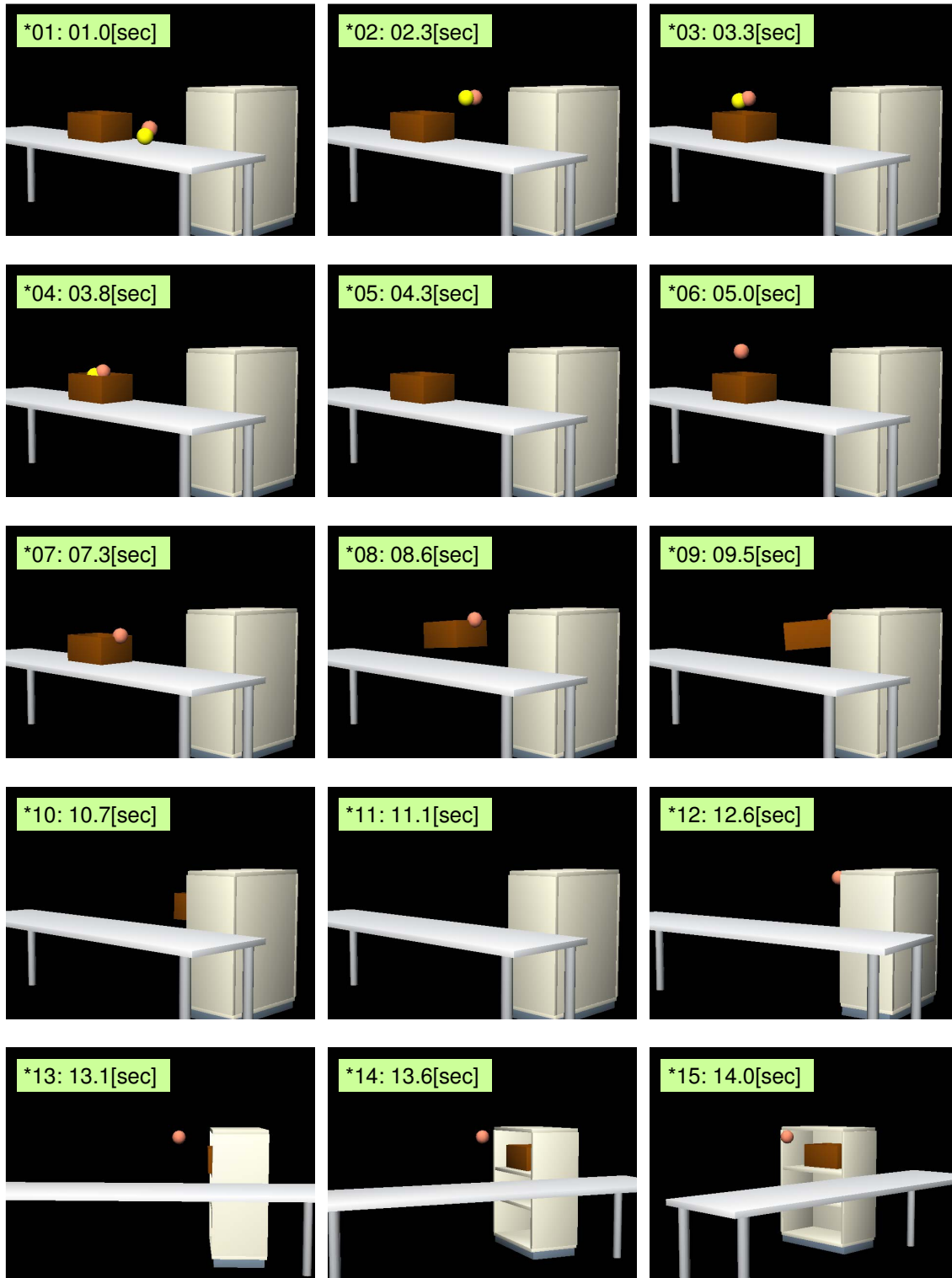


図 5.7: 箱を使って片付ける観察画像列 (著者論文 [100] の図を改変)

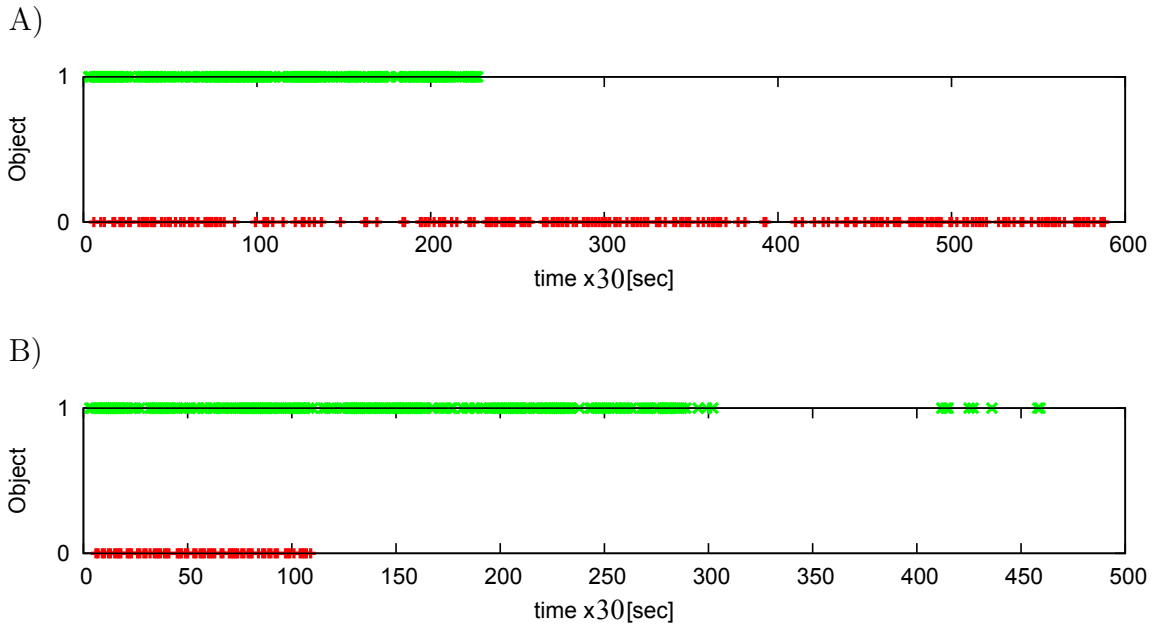


図 5.8: 探索ベース追跡手法のシミュレーションで対象が検出されたフレーム . A) が画像列 SIM-1 , B) が画像列 SIM-2 を表す (著者論文 [100] より)

objectID	time[sec]	x[mm]	y[mm]	z[mm]
...				
0	22.00	-144.5	598.6	736.3
0	22.07	-145	610.7	739.7
0	25.6	-148.4	646.4	752.5
0	25.67	-148.4	646.4	752.5
...				
...				
1	21.40	-124.3	628.3	513.3
1	23.6	-128.5	587.6	513.5
1	25.76	-127.7	594.9	512.1
...				

図 5.9: 作用推定のための入力データ列

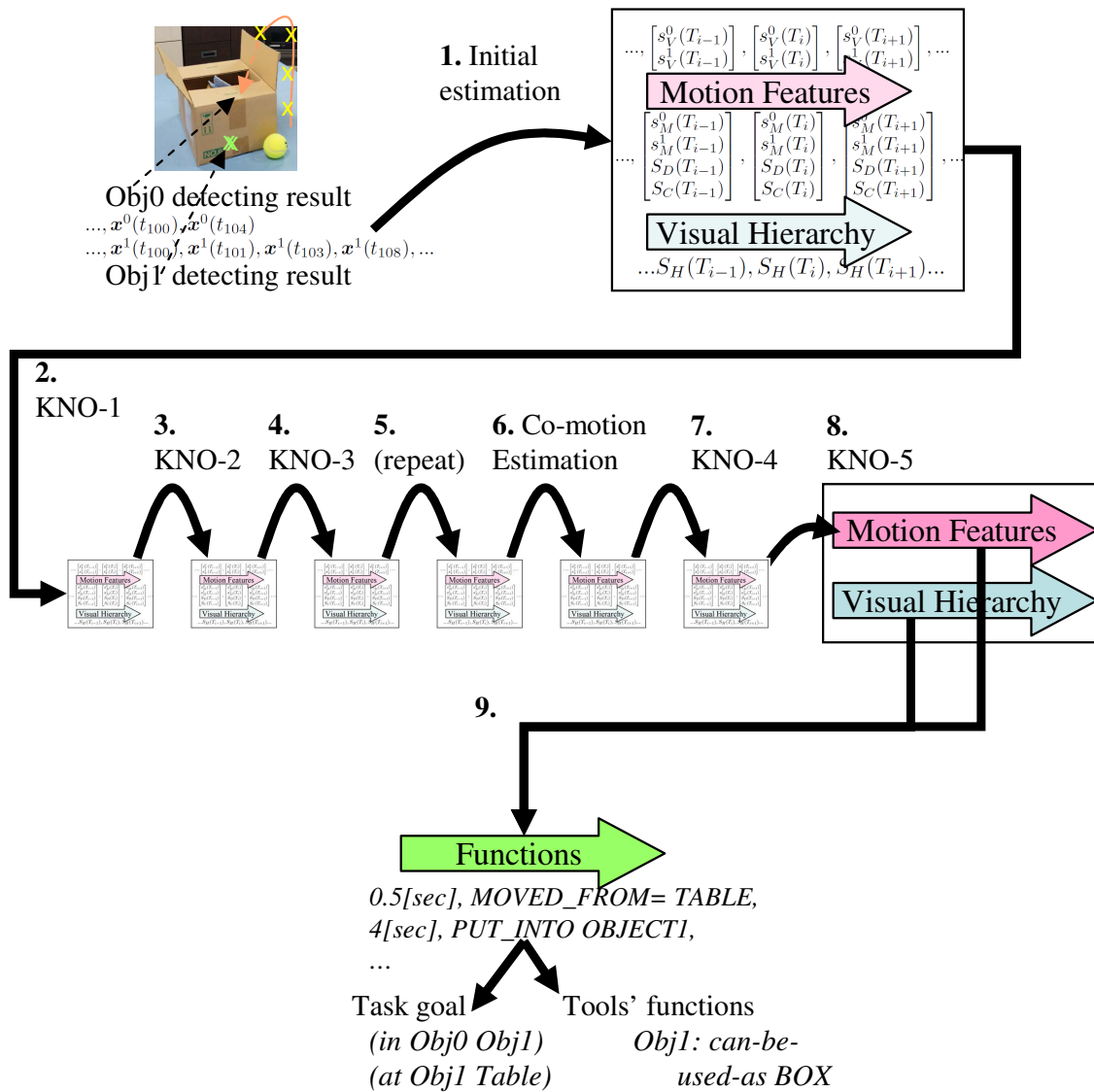


図 5.10: V-FES の概要

均位置は計算しない．以後，時区間 $T_i \in \mathbb{T}_{All}$ を単に時刻と呼び，連続する時刻集合を時区間と呼ぶものとする．

次に，各時区間における各対象の運動の有無を計算する．時刻 T_i の対象 k の速度 $\boldsymbol{v}^k(T_i)$ (mm/ t_C [sec]) は，

$$\boldsymbol{v}^k(T_i) = \bar{\boldsymbol{x}}^k(T_i) - \bar{\boldsymbol{x}}^k(T_{i-1}) \quad (5.2)$$

として求まる．この速度について閾値計算をすることで，運動の有無 $s_M^k(T_i)$ を評価する．ただし， $\bar{\boldsymbol{x}}^k(T_{i-1})$ と $\bar{\boldsymbol{x}}^k(T_i)$ が計算されていない場合，つまり $s_V^k = \text{“Unrec.”}$ の場合は “Unknown” とする．

$$s_M^k(T_i) = \begin{cases} \text{“Unknown”}, & \text{if } s_V^k(T_{i-1}) = \text{“Unrec.”} \\ & \text{or } s_V^k(T_i) = \text{“Unrec.”} \\ \text{“Yes”}, & \text{if } \|\boldsymbol{v}^k(T_i)\| \geq v_{movep} \\ \text{“No”}, & \text{otherwise} \end{cases} \quad (5.3)$$

ただし v_{movep} は，運動の有無を評価するための定数とする．

続いて，観察対象同士の距離の推移を計算する．時刻 T_i において，二つの対象が各々の位置 $\bar{\boldsymbol{x}}^0(T_i)$ ， $\bar{\boldsymbol{x}}^1(T_i)$ に認識されているとき，距離の評価値 $S_D(T_i)$ は以下のように計算される．

$$S_D(T_i) = \begin{cases} \text{“Unknown”}, & \text{if } s^0(T_i) = \text{“Unrec.”} \\ & \text{or } s^1(T_i) = \text{“Unrec.”} \\ \text{“Near”}, & \text{if } \|\bar{\boldsymbol{x}}^0(T_i) - \bar{\boldsymbol{x}}^1(T_i)\| < D_{nearp} \\ \text{“Far”}, & \text{otherwise} \end{cases} \quad (5.4)$$

ただし D_{nearp} は，距離の評価のための正の定数とする．

さらに，観察対象同士の視覚重畳関係を評価する．常に対象領域が一部でも認識可能であれば，重なり領域を推定し，画像と対象の表面のテクスチャを比較することで，上下関係を推定することが可能である [53]．しかし本章では対象が認識できない状況，対象がカメラに映らない状況をも扱う．そこで，図 5.10 右上に示すような対象の三次元軌道と，各時刻での認識の可否 ($s_V^k(T_i)$) のみから，視覚重畳関係を推定する手法を検討する．二つの対象がいずれも認識されており，かつ十分に離れている場合には，二つの視覚重畳関係は “Indp.” と決定できる．

$$S_H(T_i) = \begin{cases} \text{“Indp.”}, & \text{if } s_V^0(T_i) = s_V^1(T_i) = \text{“Yes”} \\ & \text{and } S_D(T_i) = \text{“Far”} \\ \text{“Unknown”}, & \text{otherwise} \end{cases} \quad (5.5)$$

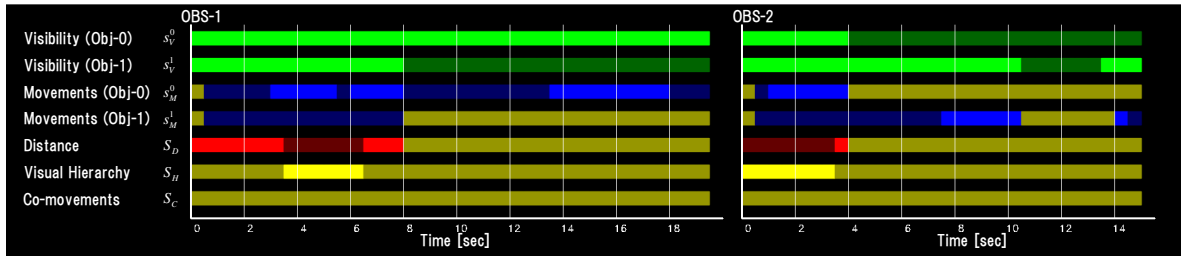


図 5.11: 疎な対象認識結果列からの物体運動，物体同士の距離，視覚重畳関係と運動随伴性の推定結果．知識を利用する前の状態である（著者論文 [100] の図を改変）

この段階で決定できるメタ情報は以上である．

SIM-OBS1 および SIM-OBS2 に，これまでに述べた手法を適用した結果を図 5.11 に示す．評価結果とグラフに描画された色との対応は図 5.5 に示す通りとし，図中 OBS1 にはコップを Obj-0，皿を Obj-1 として SIM-OBS1 の結果を，図中 OBS2 にはボールを Obj-0，箱を Obj-1 として SIM-OBS2 の結果を表した．ただし，運動の有無の閾値を $v_{movep} = 50$ [mm/unit time] とし，近接性の閾値を $D_{nearp} = 180$ [mm] としている．図 5.11 より，観察対象が見えていない区間が多く，遮蔽下で何がおこなわれているか推定する必要性が分かる．

5.3.4 知識 KNO-1 から KNO-3 の利用

一つ目の知識 KNO-1 は，過去の運動特徴と過去の視覚重畳関係から，現在の視覚重畳関係を推定するための知識である．“KNO-1: 対象同士が近づいた後に一方の対象しか見えなくなったら，その後にもう一方が見えるまでは片側遮蔽関係だろう”とする．KNO-1 が成り立つための前提条件は，対象が重要な見え隠れをする近くの時刻で，必ず対象が認識可能であることである．観測者から対象が観測可能であることが必要となることから，教示者は対象同士に重要な遮蔽関係が発生していないときに観察者へ対象を見せるようにすれば良い．

KNO-1 を適用するアルゴリズムは，次のようである．ただし，各時区間 \mathbb{T} の始めの時刻と終わりの時刻を，それぞれ $T_{init}(\mathbb{T})$ ， $T_{fin}(\mathbb{T})$ とし， $T_{i+1} = T_i + \Delta T$ のように書くものとする．

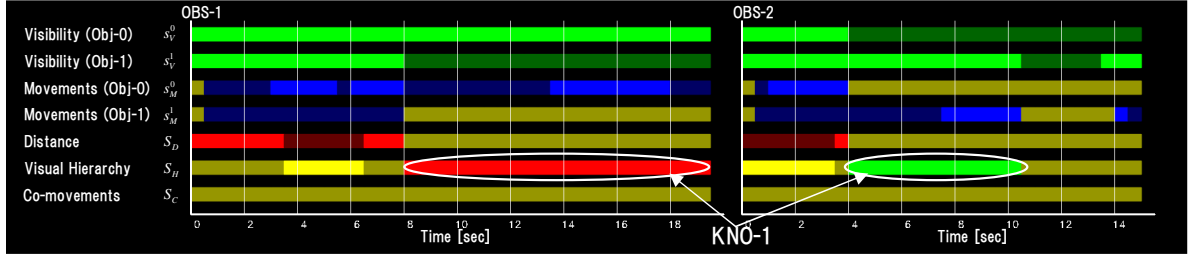


図 5.12: 知識 KNO-1 を適用して修正した後のメタ情報推移 (著者論文 [100] の図を改変)

1. 書き換え対象時区間 $\mathbb{T}_{K1,i}$ を探索する . ただし

$$\begin{aligned} \mathbb{T}_{K1,i} &\subset \mathbb{T}_{All}, S_D(T_{init}(\mathbb{T}_{K1,i}) - \Delta T) = \text{“Near”}, \\ s_V^0(T_{init}(\mathbb{T}_{K1,i})) &= \dots = s_V^0(T_{fin}(\mathbb{T}_{K1,i})) = \text{“Rec.”}, \\ s_V^1(T_{init}(\mathbb{T}_{K1,i})) &= \dots = s_V^1(T_{fin}(\mathbb{T}_{K1,i})) = \text{“Unrec.”}. \end{aligned}$$

2. 各 $\mathbb{T}_{K1,i}$ で , $S_H(T_n) = \text{“Unknown”}$ である $T_n \in \mathbb{T}_{K1,i}$ があれば , $S_H(T_n) = \text{“One-way (Obj-0)”}$ と書き換える .
3. 対象 0 と対象 1 を入れ替えて , 1., 2. をおこなう .

知識 KNO-1 を適用して修正をおこなった後の , SIM-OBS1 および SIM-OBS2 のメタ情報推移を図 5.12 に示す .

二つ目の知識 KNO-2 は , 過去の視覚重畳関係から , 現在の視覚重畳関係を推定するための知識である . “KNO-2: 片側遮蔽の後に観察対象が両方見えなくなったら , 手前の観察対象の認識失敗だろう” とする . KNO-2 が成立するための前提条件は , タスクに重要な相互作用をしよう対象は , 全てロボットが検出可能であることと , 対象が重要な見え隠れをする近くの時刻で , 必ず対象が認識可能であることである . 単なる観察対象の認識失敗であれば観察対象同士の視覚重畳関係は変化しないとすれば , KNO-2 は次のように利用できる .

1. 書き換え対象の時刻 $T_{K2,i} \in \mathbb{T}_{K2}$ を探索する . ただし ,

$$\begin{aligned} \mathbb{T}_{K2} &= \{T_{K2,i} \mid T_{K2,i} \subset \mathbb{T}_{All}, S_H(T_{K2,i} - \Delta T) = \text{“One-way”}, \\ &\quad s_V^0(T_{K2,i}) = s_V^1(T_{K2,i}) = \text{“Unrec.”}\} \end{aligned}$$

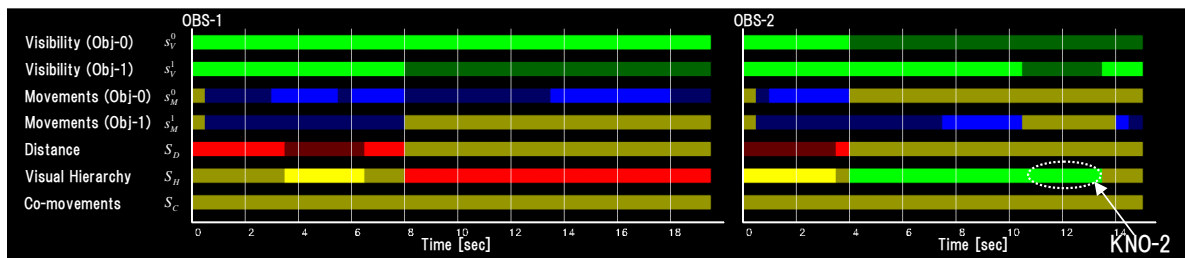


図 5.13: 知識 KNO-2 を適用して修正した後のメタ情報推移（著者論文 [100] の図を改変）

2. $T_{K2,i} \in \mathbb{T}_{K2}$ について, $S_H(T_{K2,i}) = S_H(T_{K2,i} - \Delta T)$ として書き換える.

知識 KNO-2 を適用して修正をおこなった後の, SIM-OBS1 および SIM-OBS2 のメタ情報推移を図 5.13 に示す.

三つ目の知識 KNO-3 は, 視覚重畳関係から, 運動特徴のうち近接関係を推定するための知識である. “KNO-3: 近接した後に片側遮蔽関係が発生しているということは, 物体が近いということ” とする. KNO-3 が成立するための前提条件は, 二つの対象と観測点が一直線上に並んだ状態で, 対象同士の距離が有意に変化しないことである. 観測点と二つの対象の位置関係に関する制約であることから, 教示者が観察者であるロボットへ, 対象が極端に見えにくくならないように注意してやれば良い.

KNO-3 を適用するアルゴリズムは, 次のようである.

1. 書き換え対象の時刻 $T_{K3,i} \in \mathbb{T}_{K3}$ を探索する.

$$\mathbb{T}_{K3} = \{T_{K3,i} \mid T_{K3,i} \subset \mathbb{T}_{All}, \\ S_D(T_{K3,i} - \Delta T) = \text{“Near”}, S_H(T_{K3,i}) = \text{“One-way”}\}$$

2. 各 $T_{K3,i} \in \mathbb{T}_{K3}$ で, $S_D(T_{K3,i}) = \text{“Unknown”}$ であれば, $S_D(T_{K3,i}) = \text{“Near”}$ と書き換える.

知識 KNO-2 を適用して修正をおこなった後の, SIM-OBS1 および SIM-OBS2 のメタ情報推移を図 5.14 に示す.

知識 KNO-1 と KNO-2 は過去の状態を利用するが, KNO-3 は現在の状態を利用するため, 書き換えは一度では終了しない. 書き換えが完全に終了するまで, KNO-1

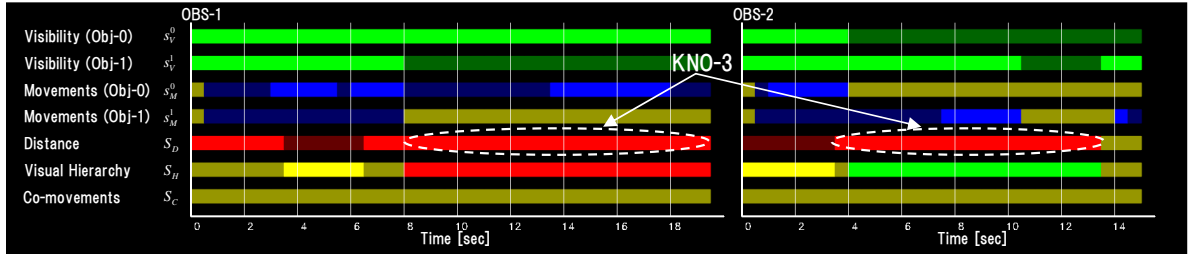


図 5.14: 知識 KNO-3 を適用して修正した後のメタ情報推移 (著者論文 [100] の図を改変)

から KNO-3 までの知識を繰り返し利用するものとする．繰り返し適用するためのアルゴリズムは次のようである．

1. フラグ `fixedp` を `fixedp = false` とセットする．
2. KNO-1 を用いたアルゴリズムを適用し，書き換わる視覚重畳関係 $S_H(T_j)$ が存在する場合のみ，`fixedp = true` とする．
3. KNO-2 を用いたアルゴリズムを適用し，書き換わる視覚重畳関係 $S_H(T_j)$ が存在する場合のみ，`fixedp = true` とする．
4. KNO-3 を用いたアルゴリズムを適用し，書き換わる近接関係 $S_D(T_i)$ が存在する場合のみ，`fixedp = true` とする．
5. `fixedp = true` であった場合は，本アルゴリズムの 1. へ戻って，もう一度計算する．`fixedp = false` である場合，計算を終了する．

知識 KNO-1 から KNO-3 を繰り返し適用して修正をおこなった後の，SIM-OBS1 および SIM-OBS2 のメタ情報推移を図 5.15 に示す．

5.3.5 運動随伴性の評価法

続いて，二物体の運動随伴性の評価をおこなう．いずれの物体も速度が計算できる場合には，二物体の速度を用いて運動随伴性を評価する [53]．時刻 T_i での運動

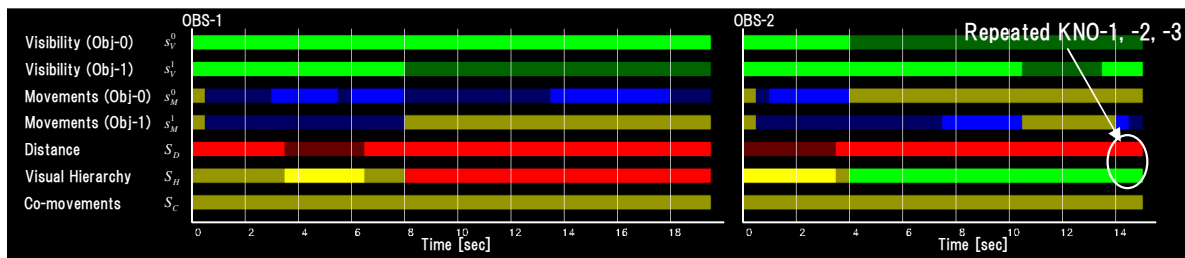


図 5.15: 観察 SIM-OBS1, 観察 SIM-OBS2 の結果へ, 知識 KNO-1 から KNO-3 を繰り返し適用した後のメタ情報推移 (著者論文 [100] の図を改変)

の随伴性評価 $S_C(T_i)$ は, 以下のように求められる.

$$S_C(T_i) = \begin{cases} \text{“Unknown”}, & \text{if } \neg h_M(T_i, 0) \text{ and } \neg h_M(T_i, 1) \\ \text{“No”}, & \text{if } (h_M(T_i, 0) \text{ and } h_M(T_i, 1) \text{ and } \neg h_c(T_i)) \\ & \text{or } (h_M(T_i, 0) \text{ and } \neg h_M(T_i, 1)) \\ & \text{or } (\neg h_M(T_i, 0) \text{ and } h_M(T_i, 1)) \\ \text{“Yes”}, & \text{otherwise} \end{cases}$$

$$h_M(t, k) \triangleq \|\mathbf{v}^k(t)\| \geq v_{thr}, h_c(t) \triangleq \|\mathbf{v}^0(t) - \mathbf{v}^1(t)\| \geq E_a$$

ただし v_{thr} と E_a はそれぞれ, 運動の有無と随伴性の有無を評価するための正の定数とする. しかしこの評価は, 物体の速度が計算できている場合のみ可能である. したがって物体が速度が計算されていない場合, つまり $s_v^k(T_i) = \text{“Unrec.”}$ あるいは $s_v^k(T_{i-1}) = \text{“Unrec.”}$ の場合には, 評価できない. この場合に対応する知識が KNO-5 であるが, この段階では計算を保留し, $S_C(T_i) = \text{“Unknown”}$ とするものとする. KNO-5 の詳細については, 続いて述べる.

SIM-OBS1 および SIM-OBS2 の観察結果に対し, 知識 KNO-1 から KNO-3 までを用いた修正を適用した後に, 運動随伴性評価をおこなった結果を図 5.16 に示す.

5.3.6 知識 KNO-4 と KNO-5 の利用

四つ目の知識 KNO-4 は, 未来の運動特徴から, 現在の運動特徴を推定するための知識である. “KNO-4: 他の対象に隠されていない対象を見失ってから再度発見し

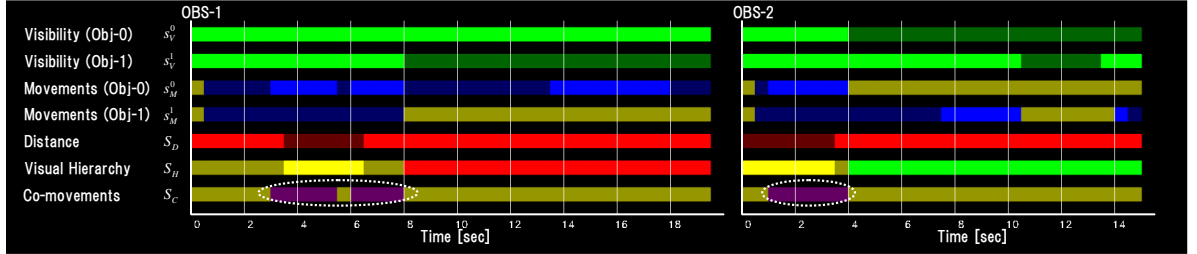


図 5.16: 運動随伴性評価をおこなった後のメタ情報推移 (著者論文 [100] の図を改変)

た場合には, 単なる認識失敗である. 位置が変わっていなければ「止まっている」, そうでなければ「動いている」とする. KNO-4 が成立するための前提条件は, 対象が見えていない間に, タスクに関連する重要な動作 (停止含む) がおこなわれていないことである. ただし物体の運搬のような, 途中での停止が重要でないタスクに関しては, 見えていない区間での対象の停止を観測できなくても問題は生じない.

KNO-4 を適用するアルゴリズムは, 次のようである.

1. 書き換え対象の時区間 $\mathbb{T}_{K4,i}$ を探索する. ただし

$$\begin{aligned} \mathbb{T}_{K4,i} &\subset \mathbb{T}_{All}, \\ s_V^0(T_{init}(\mathbb{T}_{K4,i})) &= \dots = s_V^0(T_{fin}(\mathbb{T}_{K4,i})) = \text{“Unrec.”}, \\ S_H(T_{init}(\mathbb{T}_{K4,i}), \dots, S_H(T_{fin}(\mathbb{T}_{K4,i}))) &\neq \text{“One-way (Obj-0)”}. \end{aligned}$$

2. 各時区間 $\mathbb{T}_{K4,i}$ において, $\|\bar{x}^0(T_{init}(\mathbb{T}_{K4,i})) - \bar{x}^0(T_{fin}(\mathbb{T}_{K4,i}))\| > A$ であれば, $T_n \in \mathbb{T}_{K4,i}$ について $s_M^0(T_n) = \text{“Yes”}$, そうでなければ $s_M^0(T_n) = \text{“No”}$ とする. ただし, A は正の定数とする.

3. 対象 0 と対象 1 を入れ替えて, 手順 1., 2. をおこなう.

知識 KNO-4 を用いて修正をおこなった後の, SIM-OBS1 および SIM-OBS2 のメタ情報推移を図 5.17 に示す.

五つ目の知識 KNO-5 は, 未来の視覚重畳関係と未来の運動特徴から, 現在の運動特徴を振り返り推定するための知識である. “KNO-5: 物体 X が物体 Y を隠している場合, 物体 X が十分動いたときにも片側遮蔽であれば, 物体 Y は物体 X に随伴

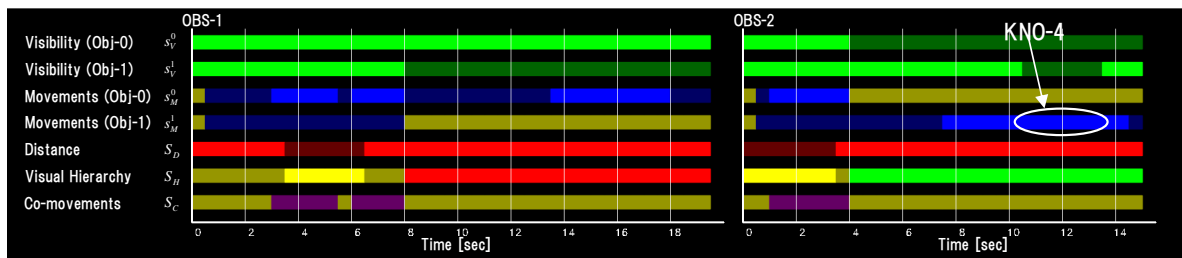


図 5.17: 知識 KNO-4 を適用して修正した後のメタ情報推移（著者論文 [100] の図を改変）

する運動をしてきた”とする．KNO-5 が成立するための前提条件は，対象が重要な見え隠れをする近くの時刻で認識可能であること，そして観察対象が消滅しないことである．

KNO-5 を適用するアルゴリズムは，次のようである．

1. 片側遮蔽の時区間 $\mathbb{T}_{K5A,i} \in \mathbb{T}_{All}$ を探索する．つまり

$$\begin{aligned} S_H(T_{init}(\mathbb{T}_{K5A,i})) &= \dots = S_H(T_{fin}(\mathbb{T}_{K5A,i})) \\ &= \text{“One-way (Obj-0)”} \end{aligned}$$

2. 各 $\mathbb{T}_{K5A,i}$ について，対象 0 が始めの場所 $\bar{x}^0(T_{init}(\mathbb{T}_{K5A,i}))$ から十分動いたと初めて分かる時刻 T_{N0} を探索する．ただし， $s_M^0(T_{N0}) = \text{“Yes”}$ ， $\|\bar{x}^0(T_{N0}) - \bar{x}^0(T_{init}(\mathbb{T}_{K5A,i}))\| > B$ とし， B は正の定数とする．さらにその後，対象 0 が動きを停止する ($s_M^0(T_{N1}) = \text{“No”}$) 時刻 T_{N1} を探索する．
3. 2. の $T_{init}(\mathbb{T}_{K5A,i})$ を T_{N1} に代えて同様に探索し，対象 0 が対象 1 を隠しながら動いている時区間 $\mathbb{T}_{K5B,i} = \{[T_{init}(\mathbb{T}_{K5A,i}), N0], [N0, N1], \dots\}$ を決定する．
4. 3. までで求めた時区間 $\mathbb{T}_{K5B,i}$ について，運動随伴性と対象 1 の運動情報の書き換えをおこなう．つまり $S_C(T_B) = \text{“Yes”}$ ， $s_M^1(T_B) = \text{“Yes”}$ ($T_B \in \mathbb{T}_{K5B,i}$) とする．
5. 対象 0 と対象 1 を入れ替えて，手順 1. から 4. までをおこなう．

知識 KNO-5 を用いて修正をおこなった後の，SIM-OBS1 および SIM-OBS2 のメタ情報推移を図 5.18 に示す．

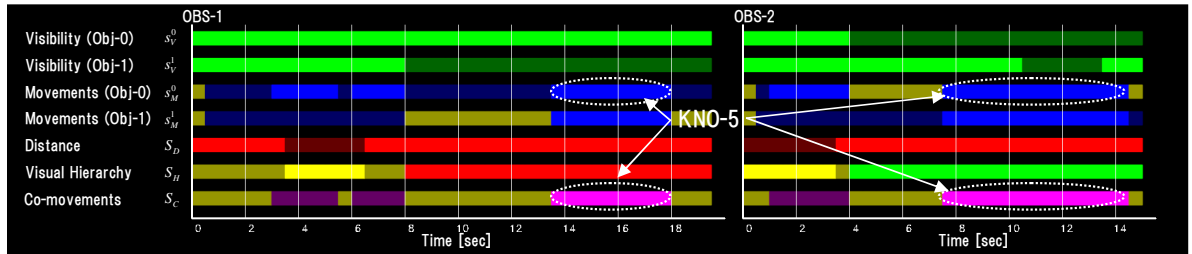


図 5.18: 知識 KNO-5 を適用して修正した後の, 最終的なメタ情報推移 (著者論文 [100] の図を改変)

5.4 時系列メタ情報からの作用・機能推定

本節では, 前節までに計算されたメタ情報を用い, 観察対象へ, いつ, どのような作用が及ぼされているかを計算する手法について述べる. さらに, 観察対象に様々な作用が及ぼされた後, タスク観察終了時での観察対象の状態としてタスク目的を, タスク中に観察対象へ及ぼされている作用の種類から道具としての機能を決定する.

本章 2 節で示したコンセプトによると, 観察対象同士の片側遮蔽関係と, その関係に至るまでの運動特徴を用いることで, 1) 物体を物体に載せる, 2) 物体を物体に入れる, 3) (それらの状態で) 移動する, という作用が推定できる.

ただし, この推定法は, 道具を下へ滑りこませる, あるいは道具で包んで持ち上げることが無いタスクであることを前提とするものである. 例外となる操作としては, 前者の例としてフライ返しを目玉焼きに差し込む操作, 後者の例として雑巾でゴミを包み持ち上げる操作が挙げられる. しかしながら, 例外の作業についても, 教示者がどちらの対象を掴んでいるかという情報を付加することができれば, どちらの対象がどちらの対象を支えながら移動しているかが推定可能だと考えられる. また, 前章で述べた空間遮蔽近傍情報を統合することでも, より詳細な関係推定が可能となる. この点については, 後述の実験 (図 5.29) で説明する. 前者の推定についても, 人の手と手の先の対象との視覚重畳関係を用いることで可能だと考えられるが, 考察に留める.

まず, 各対象が動きはじめる時刻集合 \mathbb{T}_{start} と, 動きが終わる時刻集合 \mathbb{T}_{stop} を探

索する．

$$\mathbb{T}_{start}^k = \{T_a \mid s_M^k(T_{a-1}) = \text{“No”}, s_M^k(T_a) = \text{“Yes”}\} \quad (5.6)$$

$$\mathbb{T}_{stop}^k = \{T_b \mid s_M^k(T_{b-1}) = \text{“Yes”}, s_M^k(T_b) = \text{“No”}\} \quad (5.7)$$

次に，何かに載せられる，あるいは何かに入れられる作用がおこなわれている時刻を探索する．何かに載せられる時刻 \mathbb{T}_{LOAD}^k は次のように，

$$\begin{aligned} \mathbb{T}_{LOAD}^k = \{T_c \mid T_c \in \mathbb{T}_{stop}^k, S_D(T_c) = \text{“Near”}, \\ S_H(T_c) = \text{“One-way (Obj-k)”}, S_C(T_{c-1}) \neq \text{“Yes”}\} \end{aligned} \quad (5.8)$$

何かに入れられる時刻集合 \mathbb{T}_{BOX}^k は次のように計算できる．

$$\begin{aligned} \mathbb{T}_{BOX}^k = \{T_d \mid T_d \in \mathbb{T}_{stop}^k, S_D(T_d) = \text{“Near”}, \\ S_H(T_d) = \text{“One-way (Obj-}\bar{k}\text{)”}, S_C(T_{d-1}) \neq \text{“Yes”}\} \end{aligned} \quad (5.9)$$

ただし，対象 \bar{k} は対象 k でない対象を指すものとする．

続いて，対象が他の対象に隠れながら動いている時区間を推定する．動き始める時刻集合 \mathbb{T}_{start*} と，動き終わる時刻集合 \mathbb{T}_{stop*} は，以下のように計算できる．

$$\begin{aligned} \mathbb{T}_{start*}^k \{T_e \mid T_e \in \mathbb{T}_{start}^k, S_D(T_e) = \text{“Near”}, \\ S_H(T_e) = \text{“One-way (Obj-k)”}, S_C(T_{e+1}) = \text{“Yes”}\} \\ \mathbb{T}_{stop*}^k \{T_f \mid T_f \in \mathbb{T}_{stop}^k, S_D(T_f) = \text{“Near”}, \\ S_H(T_f) = \text{“One-way (Obj-k)”}, S_C(T_{f-1}) = \text{“Yes”}\} \end{aligned} \quad (5.10)$$

ただし， \mathbb{T}_{start*} あるいは \mathbb{T}_{stop*} に含まれる時刻では，他の対象に隠されている対象の場所が分からない．しかし，隠している方の対象では場所が分かることを利用すれば，以下のように場所の伝達をおこなうことができる．

$$\begin{aligned} \mathbf{x}^k(T_n) = \mathbf{x}^{\bar{k}}(T_n), \text{ if } s_V^k(T_n) = \text{“Yes”} \\ (T_n \in \mathbb{T}_{start*}^k \text{ or } T_n \in \mathbb{T}_{stop*}^k, k = 0, 1) \end{aligned} \quad (5.11)$$

さらに，各対象の動き始めの時刻，あるいは止まる時刻について，それがロボットが予め知っている場所であったならば，その場所を表す記号で置き換える（スナップする）．これはシンボリックなタスクプランニングを可能とするためである．世界座標系での観察対象の位置を直接記憶しシンボルとして利用しても良いが，その場

合はタスクを複数回見せられた場合に、場所のわずかな違いにより違うシンボルとなってしまう、学習結果を統合することが難しいと考えられる。実際の家事タスクの目的では、物が机の上に置かれること、箱の中に入ることなど、動かす先の物体やその面が重要となる場合が多いであろうことから、場所を表す記号は、幅を持った表現であるものとした。具体的には、机の代表点からの距離が一定値以内のものを机とするなどとする。後述の実験では、観察対象が動かされる先の家具は既知であるものとしたが、未知の場所が含まれる場合には、家具あるいは平面認識等をおこない、その場所を覚えるといった作業をおこなえば良い。ただし、このような場所のシンボル化をおこなう場合は、シンボルの幅が学習内容の粒度を決定する。机の右側に置くことが重要、などといった場合を扱うためには、ロボットが目的地の推定結果をユーザに示し、ユーザがインタラクティブに目的となる範囲を指示・変更していく仕組みが必要である。

また、物体の移動が主な目的である場合、空中での一時停止は省略して良いと考えられるため、後述の実験でも略すものとした。さらに、ある場所から同じ場所へ移動するという、繰り返しの動きも消去する。物体の移動そのものが目的ではない、物を振り混ぜる、雑巾で机を拭くなどの作業においては、繰り返しの動きは重要となる。このような操作が含まれる一連の作業の学習においては、観察対象の軌道をより詳細に観察していく必要があるが、それほど遮蔽が問題とならない作業であれば、軌道を得ることは大きな問題にはならないと考えられる。

最後に、各対象の道具としての機能を推定する。 T_{LOAD}^k , T_{BOX}^k の時刻が存在する場合、それぞれ対象 \bar{k} には上にものを載せることのできる機能“MAT”，中にものを入れることのできる箱のような機能“STORAGE”が存在すると推定する。さらにその後、動きが観測された場合には、 \bar{k} には上にものを載せて運ぶことのできる機能“TRAY”，中にものをに入れて運ぶことのできる機能“CONTAINER”があると推定する。

このように計算される道具的機能とタスク目的は、シンボリックなタスクプランニングへ利用可能である [104]。詳細は次章で述べる。

本節で述べたタスク目的・機能推定法を適用した結果を図 5.19，図 5.20 に示す。ただし、SIM-OBS1 では机とキッチン，SIM-OBS2 では机と棚に対象がスナップ可能であるようにした。結果，SIM-OBS1 では皿にコップを重ねた状態でキッチンに

```

1 * OBJECT0
2 AT TIME=2.5[sec], MOVED_FROM PLACE= TABLE,
3 AT TIME=8[sec], LOADED_ONTO OBJECT1,
4 (AT TIME=13[sec], START being TRANSPORTED, by OBJECT1)
5 (AT TIME=18[sec], STOP being TRANSPORTED, by OBJECT1)
6 * OBJECT1
7 AT TIME=13[sec], MOVED_FROM PLACE= TABLE,
8 AT TIME=18[sec], MOVED_TO PLACE= KITCHEN,
9 ;; TOOL INFORMATION:
10 * OBJECT1 CAN BE USED AS MAT
11 * OBJECT1 CAN BE USED AS TRAY
12 ;; TASK GOAL:
13 ON OBJECT0 OBJECT1
14 AT OBJECT1 KITCHEN

```

図 5.19: SIM-OBS1 の観察学習結果

置くこと，SIM-OBS2 ではボールを箱に入れて棚に置くことがタスクの目的であると推定された．さらに，皿には上に物を重ねて良いことと重ねた状態で運んで良いこと，箱には中に物を入れて良いこととその状態で運んで良いことが正しく推定された．

5.5 ロボットによる実観察画像列を用いた観察学習実験

本節では，前節までで提案した手法を実ロボットでの観察画像列に適用し，その有用性を確認する．ロボットはHRP-2VZ (CV-Head Mk.II) を用いるものとし，色画像および距離画像カメラを用いたタスク観察をおこなう．

5.5.1 柔軟物を含むタスク観察学習法

提案手法には，観察対象がフレームアウトするなどして認識できない区間が長く存在しても問題ないという利点がある．さらに，物体の詳細な三次元形状情報や，物体同士の複雑な位置関係や接触関係を陽に観測できなくても，視覚重畳関係の変化等から入った・載せた等の作用が認識可能であるという利点もある．

```
1 * OBJECT0
2 AT TIME=0.5[sec], MOVED_FROM PLACE= TABLE,
3 AT TIME=4[sec], PUT_INTO OBJECT1,
4 (AT TIME=7[sec], START being TRANSPORTED, by OBJECT1)
5 (AT TIME=14.5[sec], STOP being TRANSPORTED, by OBJECT1)
6 * OBJECT1
7 AT TIME=7[sec], MOVED_FROM PLACE= TABLE,
8 AT TIME=14.5[sec], MOVED_TO PLACE= SHELF,
9 ;; TOOL INFORMATION
10 * OBJECT1 CAN BE USED AS STORAGE
11 * OBJECT1 CAN BE USED AS CONTAINER
12 ;; TASK GOAL:
13 IN OBJECT0 OBJECT1
14 AT OBJECT1 SHELF
```

図 5.20: SIM-OBS2 の観察学習結果

本節の実験では、このような利点を評価すべく、観察対象として柔軟物である布袋を取り上げる。布袋に光の透過性が無く、もう一つの観察対象が布袋より小さな人形（ぬいぐるみ）であった場合、二つの対象同士の関係には、

- 1) ぬいぐるみと布袋が別々に存在する、
- 2) ぬいぐるみが布袋の中に入っている
だけでなく、
- 3) ぬいぐるみが布袋の上に載っている（布袋が敷物としての役割を果たす）、
- 4) ぬいぐるみの上に布袋がかけられている（布袋がホコリよけとしての役割を果たす）

の四種類の状態が存在する。特に2)と4)の状態では、いずれもぬいぐるみは布袋に遮蔽され、視覚認識が不可能である。これらを識別し、さらに単なるフレームアウトとも分類するためには、遮蔽下での状況を推定する手法を用いる必要がある。

観察実験は二回おこない（観察 HRP2-OBS1，観察 HRP2-OBS2），ぬいぐるみを入れる，布袋をかける，フレームアウトする状況が全て入るものとする。観察 HRP2-OBS1 と観察 HRP2-OBS2 の様子を，各々図 5.21 と図 5.22 に示す。観察 HRP2-OBS1 では，図 5.21 に示す通り，ぬいぐるみを袋に入れてから（図中，2か

ら3) 棚へ運ぶ。一方，観察 HRP2-OBS2 では図 5.22 のように，まずぬいぐるみを棚へ運んでから（図中，3），布袋を運んでぬいぐるみへかける（図中，7）。

観察学習において観察対象はぬいぐるみと布袋とし，認識には各々の色を手がかりとする手法を用い，以下の手順で各時刻の三次元位置を計算する：

- 1) まずロボット頭部カメラから得られた色画像のうち，対象色の領域を抽出する。抽出には HSV 色空間を用いた。
- 2) 膨張・収縮処理をおこなってノイズ削減後，ラベリングをおこない，適切な大きさの対象色領域が存在する場合に，それを観察対象の領域とする。
- 3) 2 で観察対象が認識できている場合のみ，観察対象の領域の重心位置の距離画像の値を用いて，カメラ座標系での観察対象の三次元位置を決定する。
- 4) さらに，ロボットの関節角度値を用いて，世界座標系での観察対象の三次元位置を決定する。

この手法では，観察対象と類似する色および大きさの対象が，観察学習環境に他に存在しないことが前提条件となる。しかし条件を満たせば，本実験で対象とするような変形の起こり得る柔軟物体を対象とする場合にも，偽陽性を含まない認識結果を出力することが可能である。

この手法により，観察 HRP2-OBS1 および観察 HRP2-OBS2 において二つの観察対象を追跡した結果を，図 5.23 および図 5.24 に示す。いずれの実験においても観察対象がフレームアウトした状況，あるいはぬいぐるみが布袋に隠された状況では，対象の認識はできていない。ただし二つの追跡結果ともに，観察対象が見えていないのに見えた，あるいは観察対象以外の場所に認識された，といった偽陽性結果は含まれていない。

5.5.2 実観察画像列を用いたタスク学習結果

前小節の，二つの観察対象の追跡結果を V-FES の入力として用い，視覚重畳関係，運動特徴を推定した結果を図 5.25 および図 5.26 に示す。評価結果とグラフに描画された色との対応は図 5.5 に示す通りとする。知識を利用する前のメタ情報の時間推移が図 5.25 の (a) であるが，二つの観察状況ともに，対象のフレームアウトや遮蔽に起因した認識の失敗が長期間に表れている (s_V)。図 5.25 の (b) が知

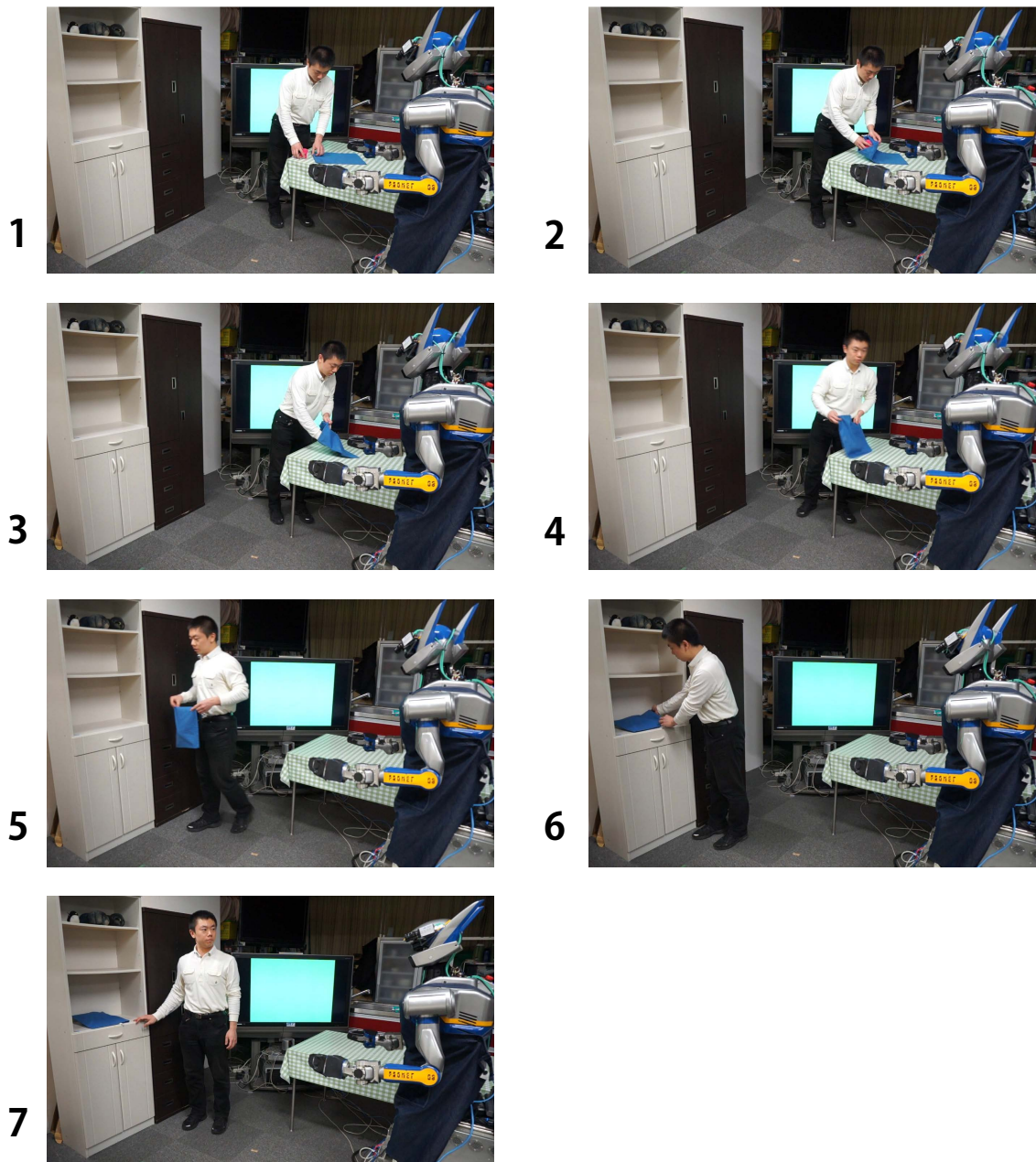


図 5.21: 観察 HRP2-OBS1 の様子

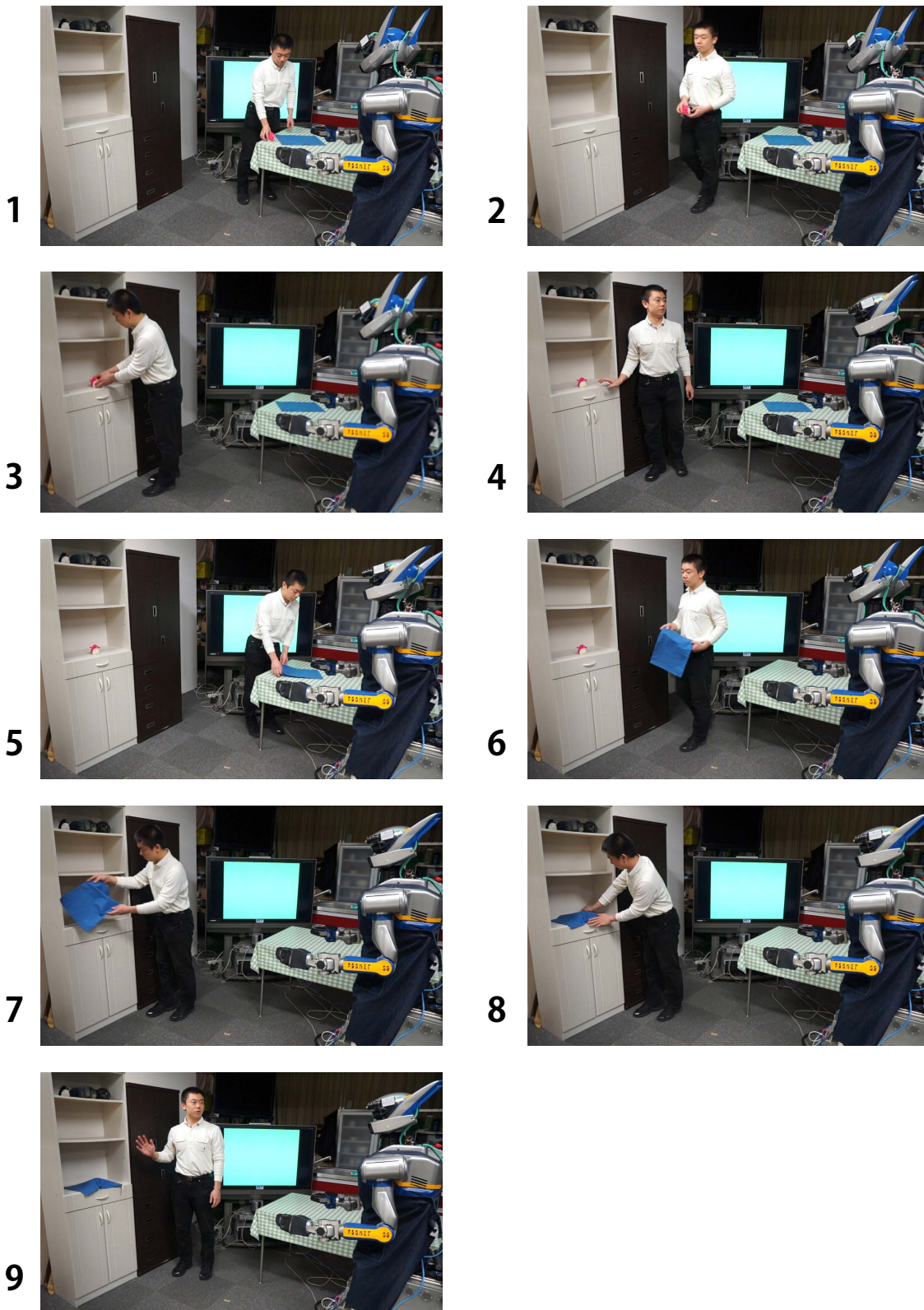


図 5.22: 観察 HRP2-OBS2 の様子

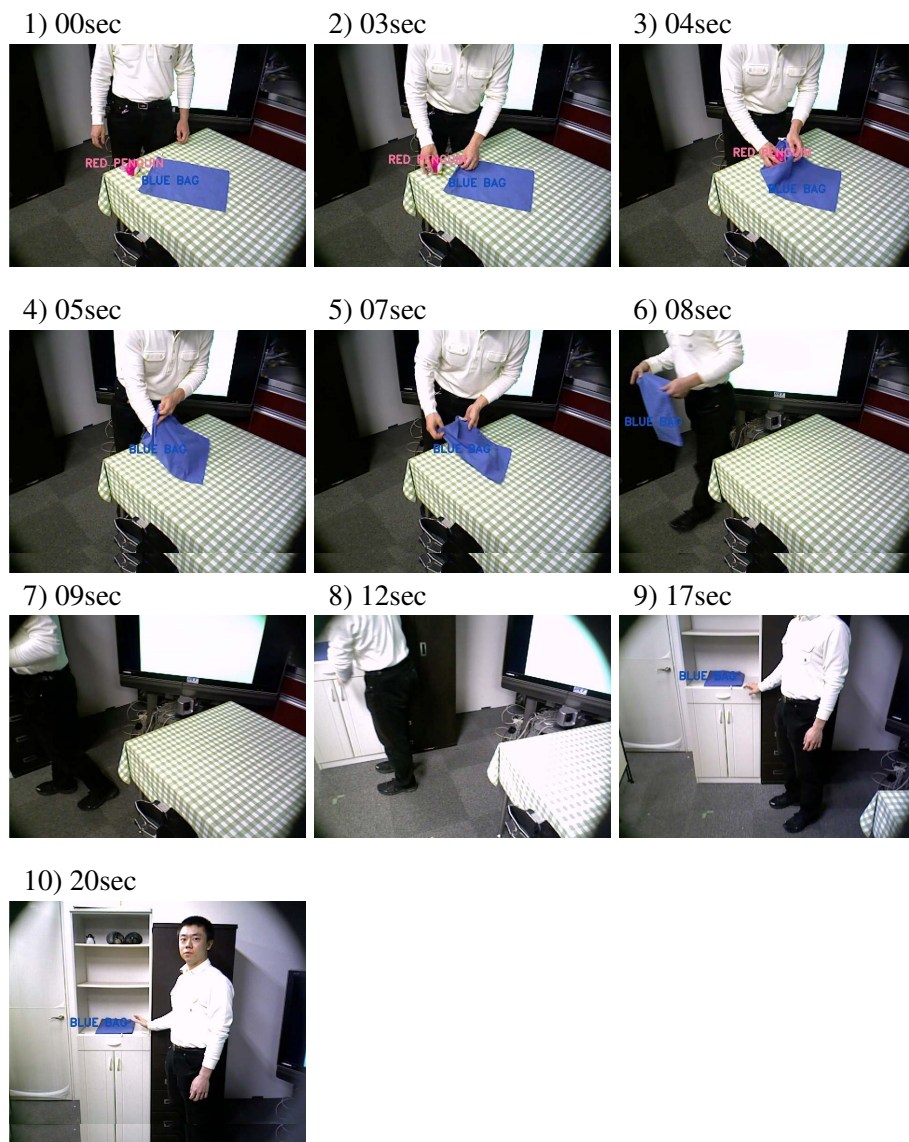


図 5.23: 観察 HRP2-OBS1 で、ぬいぐるみと青い袋を認識した結果．各々が認識された時刻には，その位置に対象の名称を重畳表示している

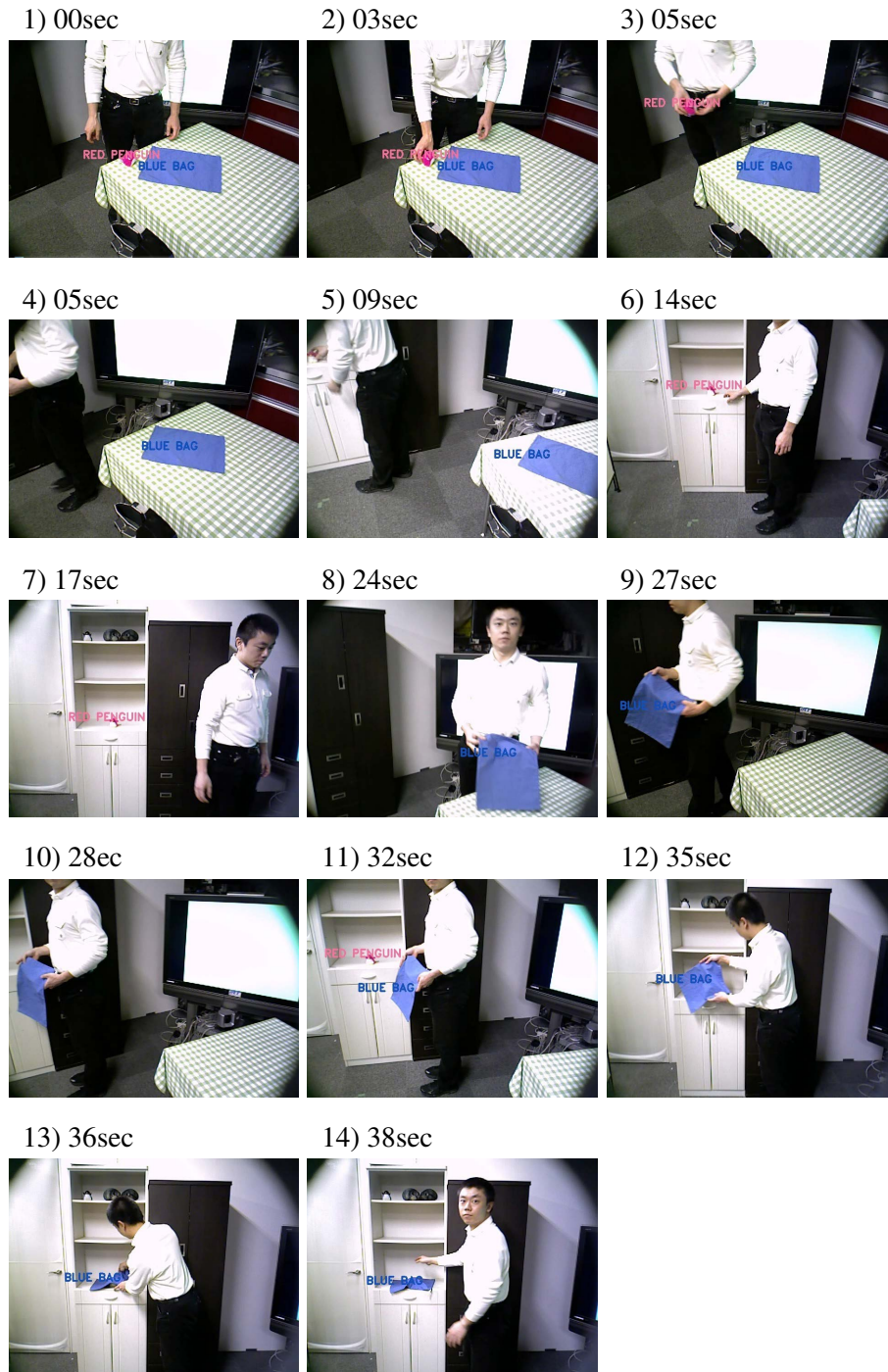


図 5.24: 観察 HRP2-OBS2 で、ぬいぐるみと青い袋を認識した結果．各々が認識された時刻には，その位置に対象の名称を重畳表示している

知識 KNO-1 から知識 KNO-3 までを繰り返し利用した結果である．観察 HRP2-OBS1 では，ぬいぐるみを袋に入れた後にもぬいぐるみと袋の距離が十分に近いこと，そしてぬいぐるみと袋に重畳関係が発生していることが正しく推定できている．観察 HRP2-OBS2 においても，ぬいぐるみに布袋をかぶせた際に，両対象の距離が近く，重畳関係が発生していることが推定されている．図 5.26 に示す最終的な時系列メタ情報の推定結果では，観察 HRP2-OBS1 においてぬいぐるみが袋に入っている間に動かされたこと，両物体がフレームアウトしている間に動かされていることが正しく推定されている．観察 HRP2-OBS2 においてもフレームアウト時の運動情報が推定できしており，V-FES の時系列メタ情報の相補的推定が正しくおこなわれていると言える．

さらに V-FES により，観察中に表れる作用と，観察対象に表れる道具としての機能を推定させた．ただし，作用推定における場所のシンボル化では，机の上および棚の上の領域を予め与えるものとした．観察 HRP2-OBS1，観察 HRP2-OBS2 それぞれの推定結果を図 5.27，図 5.28 に示す．

観察 HRP2-OBS1 では OBJECT0 (ぬいぐるみ) を OBJECT1 (袋) に入れてから運んだこと，観察 HRP2-OBS2 では OBJECT0 を運んでから OBJECT1 を載せたことが正しく推定されている．さらに観察 HRP2-OBS1 では OBJECT1 に関して，物を入れる使い方 (STORAGE) と入れた状態で運ぶ使い方 (CONTAINER) が可能であることが正しく推定され，観察 HRP2-OBS2 では，OBJECT0 の上に物を載せて良いこと (MAT) が推定されている．ぬいぐるみは敷物ではないものの，その上に物を載せて運ぶ機能 (TRAY) が無いことが推定できているため，上にもものかけることだけが可能であることが学習できていると言える．

このように作用や道具的機能を推定・学習することの利点の一つは，実験で示したように，形状が定まっていない柔軟物に関しても，入れる・載せる (かける) といった作用や機能を推定することができる点である．布地と，もう一方の観察対象の詳細な三次元形状がどちらも推定できれば，入れられ・載せる・載せられる (敷く) という作用は推定できるであろう．しかし，柔軟性のある布地の三次元形状を把握することは難しく，特に裏側の生地を知ることは，遮蔽により困難である．一方で本手法によると，布地以外の観察対象の動きや視認性から，作用の推定が可能であり，詳細な三次元形状の推定なしに布地の道具的機能を推定することが

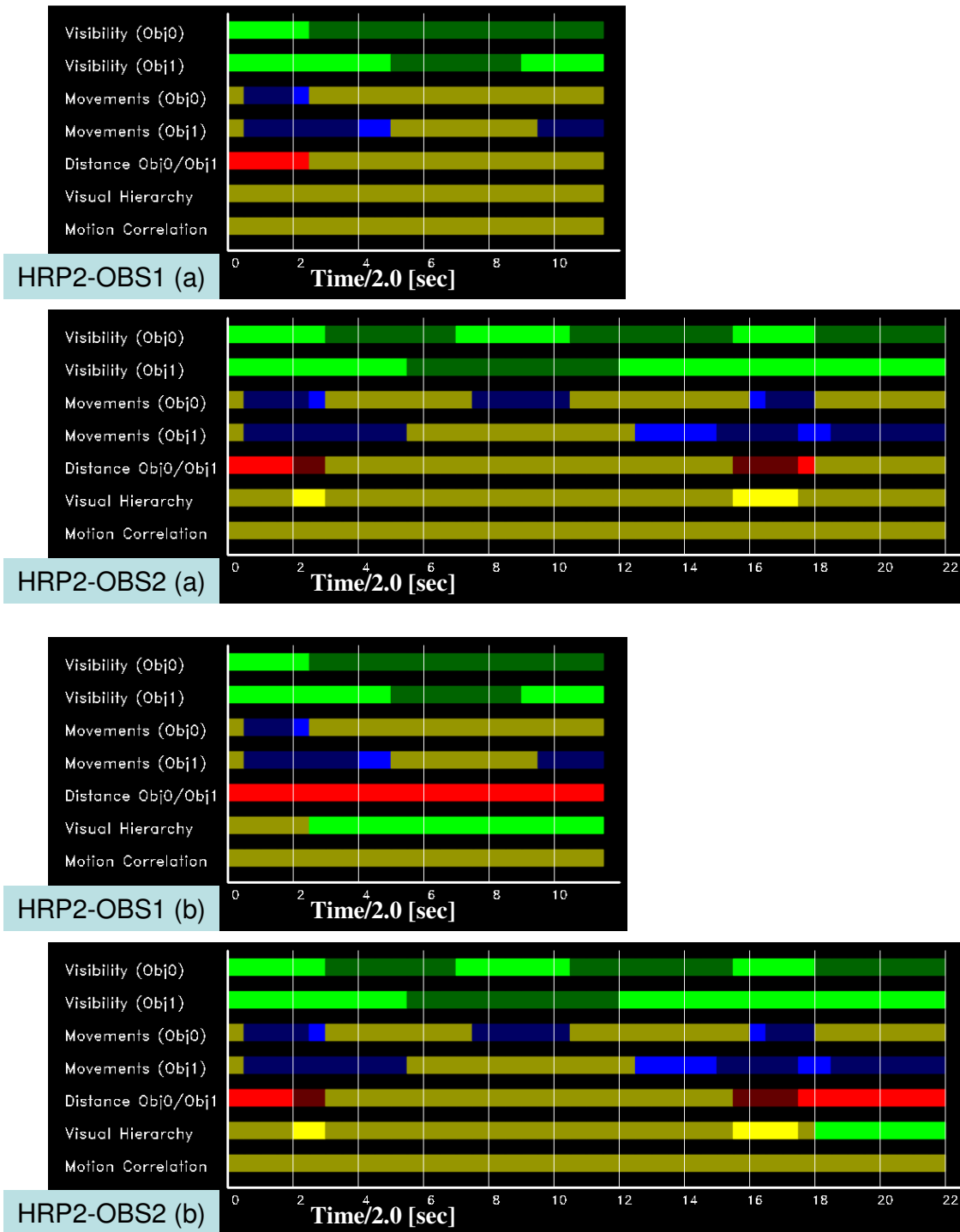


図 5.25: 観察 HRP2-OBS1 および HRP2-OBS2 における時系列メタ情報推定過程。(a) が知識を利用する前, (b) が知識 KNO-1 から知識 KNO-3 までを繰り返し利用した結果である

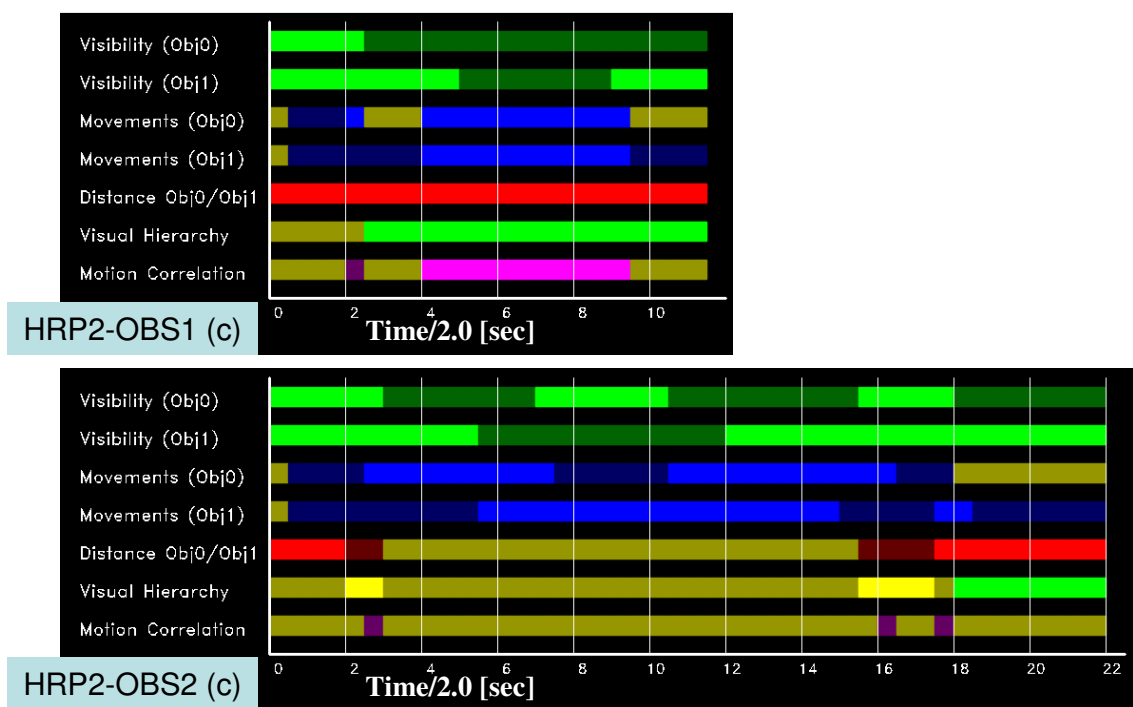


図 5.26: 観察 HRP2-OBS1 および HRP2-OBS2 における, 最終的な時系列メタ情報推定結果

```

* OBJECT0
AT TIME=1.5[sec], MOVED_FROM PLACE= TABLE,
AT TIME=2.5[sec], PUT_INTO OBJECT1,
(AT TIME=3.5[sec], START being TRANSPORTED , by OBJECT1)
(AT TIME=9[sec], STOP being TRANSPORTED , by OBJECT1)
* OBJECT1
AT TIME=3.5[sec], MOVED_FROM PLACE= TABLE,
AT TIME=9[sec], MOVED_TO PLACE= SHELF,
* OBJECT1 CAN BE USED AS STORAGE
* OBJECT1 CAN BE USED AS CONTAINER
;; TASK GOAL:
IN OBJECT0 OBJECT1
AT OBJECT1 SHELF

```

図 5.27: むいぐるみを袋に入れてから運ぶ観察 HRP2-OBS1 の作用推定・道具機能推定結果

```

* OBJECT0
AT TIME=2[sec], MOVED_FROM PLACE= TABLE,
AT TIME=7.5[sec], MOVED_TO PLACE= SHELF,
* OBJECT1
AT TIME=5[sec], MOVED_FROM PLACE= TABLE,
AT TIME=18.5[sec], LOADED_ONTO OBJECT0,
* OBJECT0 CAN BE USED AS MAT
;; TASK GOAL:
AT OBJECT0 SHELF
ON OBJECT1 OBJECT0

```

図 5.28: むいぐるみを運んでから布をかける観察 HRP2-OBS2 の作用推定・道具機能推定結果

できる．逆にこの機能情報を用いて，およその布地の形状情報を推定することも将来展望として考えられる．

ただし，本実験で使用した作用推定法では，前節で述べたように，布地に近づく観察対象の遮蔽のされ方の詳細を利用していないため，布地の下に観察対象が滑りこむような場合にも“入った”と認識されてしまう．このような誤認識を無くす一手法として，前章で用いた相互重畳関係を利用する手法が挙げられる．相互重畳関係の推定には，観察対象の領域を計算できている必要があるが，本節で用いている認識手法は，観察対象それぞれの領域をも計算するものであるため，利用可能である．

このように，遮蔽時刻近傍情報と空間近傍情報の両方を用い，ぬいぐるみと布地の関係推定をおこなった発展的な実験を図5.29に示す．この実験では，ぬいぐるみと布地の領域を本節で用いた領域推定法によって計算し，ぬいぐるみの視認性が変化する時刻の前の時刻の状態により，ぬいぐるみと布地の関係を推定している．1) ぬいぐるみと布地に重畳関係が無い状態でぬいぐるみが見えなくなった場合には，ぬいぐるみが下へ潜り込んだとし，2) ぬいぐるみが布地に被さる重畳関係の場合は，ぬいぐるみが布地の上にあるとする．さらに，3) ぬいぐるみが布地の上にある状態から見えなくなった場合には，ぬいぐるみは布地の中にあるとする．ただし，布地とぬいぐるみの重畳関係のうち片側遮蔽・独立関係の推定には，布地領域の凸包の中にぬいぐるみが含まれているか否かを用い，3) では相互重畳関係を簡易的に推定させている．布地領域に穴が無い場合には，この推定法は妥当だと考えられるが，実際にぬいぐるみと布地の“inside”，“on”，“under”の関係が推定できている．これは時空間遮蔽近傍特徴の融合の有用性を示すものだと考えられる．ただし，このように空間遮蔽近傍特徴情報を使う場合でも，観察対象同士の間関係を正しく推定するためには，観察対象を高密度に捉える必要がある．片付けのような大きな移動が伴うタスクの観察においては，3章で述べたような，広範囲の注意と高視力の視覚を統合する仕組みが必要であろう．

また，ぬいぐるみのような，人間の感覚ではマットやトレーとは異なる物体に関しても，“MAT”機能と“TRAY”機能を分けて学習することに利点が存在する．本実験の状況では，ぬいぐるみを使って布を運ぶことは通常考えられない．本手法では，ぬいぐるみに関する道具的利用法として“TRAY”機能が推定されないため，タスク再現にあたっては，ぬいぐるみを置いてから布をかけるという操作のみが選択

され、このような制約を満たすタスクが計画されることはメリットだと考えられる。

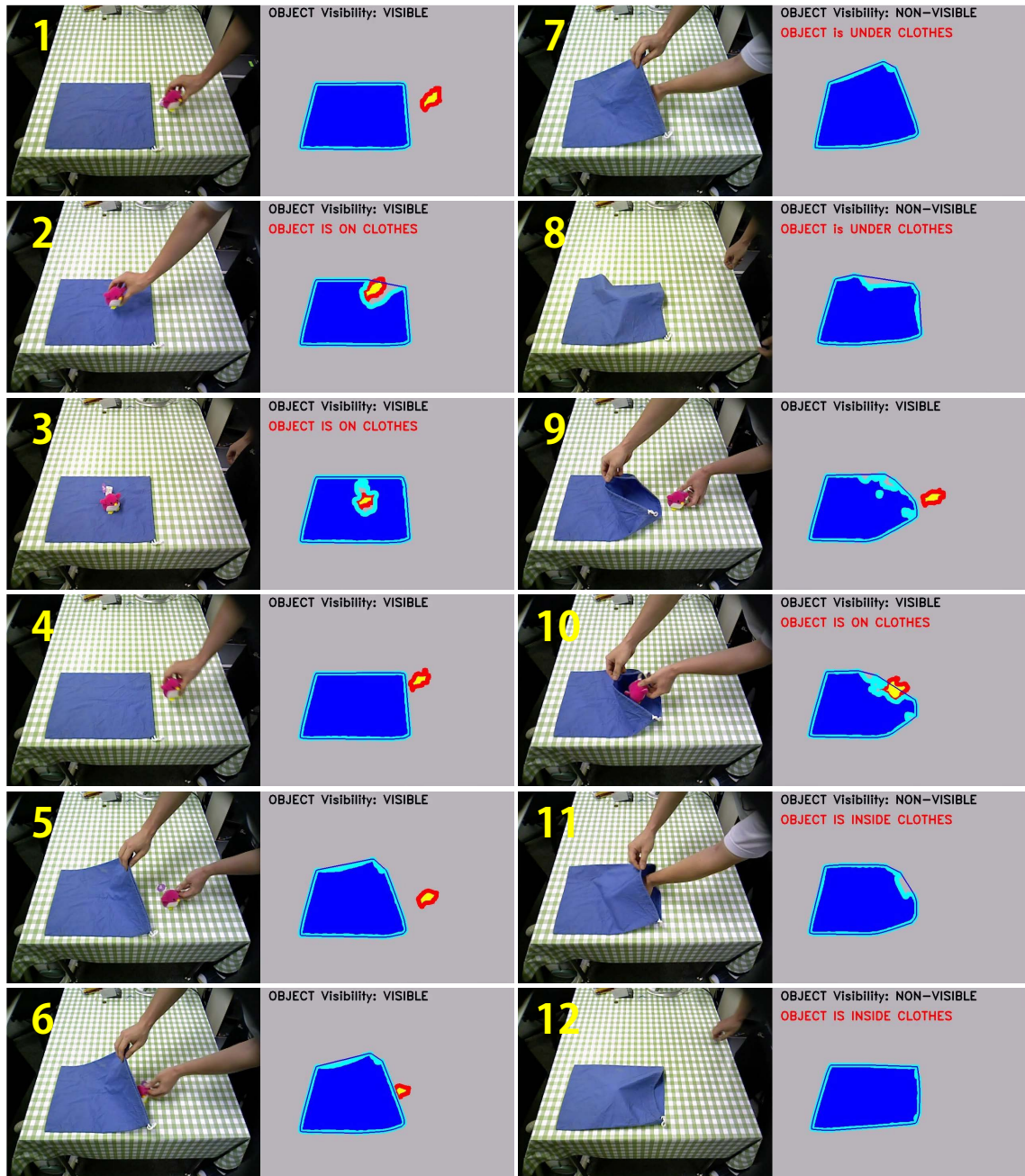


図 5.29: 時空間遮蔽近傍情報を利用した,柔軟物を含む観察対象同士の関係推定実験

5.6 LINE-MOD を用いたテクスチャレス物体追跡と観察学習

本章で提案した V-FES は，偽陽性の低い物体探索結果を出力するものであれば，任意の視覚探索手法との組み合わせが可能なよう設計されている．本節ではこの性能を評価するため，前節で用いた手法とは異なる，観察対象の三次元形状情報に基づくロバストな視覚探索手法を組み合わせ，評価する．

5.6.1 LINE-MOD

本研究で対象とする日用品では，食器など，テクスチャの少ない物体が多く存在する．テクスチャの少ない物体でも，三次元形状情報を併用することで認識可能な手法として，LINE-MOD[31] がある．LINE-MOD は，色画像のエッジ特徴に加えて，距離画像の法線方向特徴を用いてテンプレートマッチングをおこなうことで，ロバストな物体の発見および三次元座標推定をおこなう手法である．本節では，観察対象となる物体の認識に，LINE-MOD における影の影響を削減するよう変更した改良 LINE-MOD[109] を用いる手法および，その観察学習結果について評価する．

LINE-MOD・改良 LINE-MOD では，認識対象の表面の三次元点群をテンプレートの入力として用いる．テンプレート作成のための三次元モデリングの様子を，図 5.30 に示す．マーカを用いて位置合わせをしながら，認識対象を全方位から RGBD カメラで撮影し，1分から数分ほどでのモデリングが可能となっている．テンプレートマッチングの際には，記憶している三次元点群から様々な視点におけるエッジテンプレート画像と法線方向テンプレート画像を生成して用いる．入力された色画像からエッジを，距離画像から法線方向画像を計算し，エッジテンプレート画像と法線方向テンプレート画像を局所変形させながらマッチングすることで，ロバストなテンプレートマッチングを実現している．

5.6.2 LINE-MOD の物体認識特性に関する考察

まず予備実験により，物体が重なり合う状況下での，改良 LINE-MOD の認識特性について検討した．

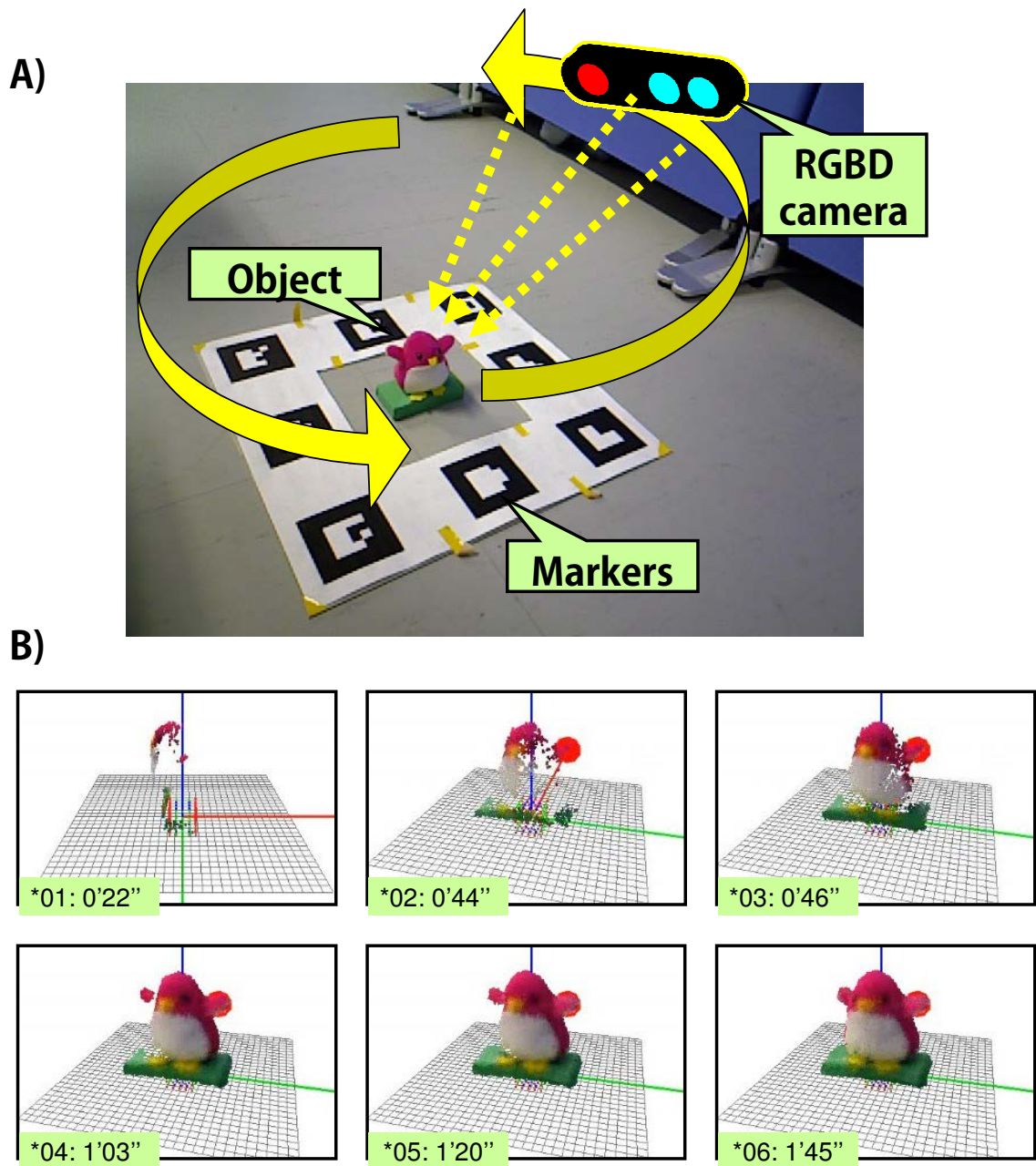


図 5.30: LINE-MOD テンプレート用の三次元点の取得 . A) は三次元点群を取得するために使用する環境である . B) は作成中のテンプレートを示す

表 5.1: 改良 LINE-MOD による物体の認識特性：遮蔽率 20 % 未満を陽性と期待する場合

	True positive	False positive	False negative	True negative	Number of trials
Cup	279	0	80	0	359
Dish	317	0	0	42	359

予備実験 1 ではコップを皿に載せた後，皿を持って動かす状況を，RGBD カメラ (ASUS Xtion PRO LIVE) で撮影した入力画像列を用い，改良 LINE-MOD でコップと皿を認識させた．実験中に対象物を検出したフレームにおいてテンプレートを重ね表示した結果，そして各物体が認識された時刻を表したグラフを，各々図 5.31 と図 5.32 に示す．全 359 回の認識ループのうち，コップと皿を検出したのは各々 279 回，317 回であった．

図 5.32 から，1) コップは全般的に検出されているが，検出に失敗している時刻も存在すること，2) 皿は前半のみ検出されているが，後半は検出失敗していることが分かる．またコップが検出された 279 フレーム，および皿が検出された 317 フレームにおいて，検出結果は全て正しく，追跡すべき対象を捉えていた．1) に関して，コップが静止あるいは動きが遅い場合には検出されているが，動きが速い場合には検出されていない．これは，改良 LINE-MOD が対象のブレには頑強でないからだと考えられる．2) に関して，改良 LINE-MOD は遮蔽に弱いことがわかる．しかし遮蔽がない場合，背景と認識対象の皿の色が近似しているにもかかわらず認識できている．結果全般より，改良 LINE-MOD は遮蔽および激しい動きには脆弱だが，その他の状況には頑強な認識手法であり，家庭環境内での観察に有用だと考えられる．

次に，改良 LINE-MOD の遮蔽に対する認識特性を評価した．認識対象物の遮蔽率が 20 % を超える場合を陰性，超えない場合を陽性と期待して偽陽性および偽陰性を計算すると，表 5.1 のようになった．ただし遮蔽率とは，色画像中の認識対象物の面積に基づき，実際に色画像中に現れている認識対象物の面積を，遮蔽物がない場合に投影されるべき面積で割ったものとする．表 5.1 によると，皿，コップともに偽陽性がゼロである¹．

さらに基礎実験二つだけでなく，後述する HSR による観察実験 HSR-OBS1，観

¹正確には，偽陽性をゼロにするパラメタを設定可能である

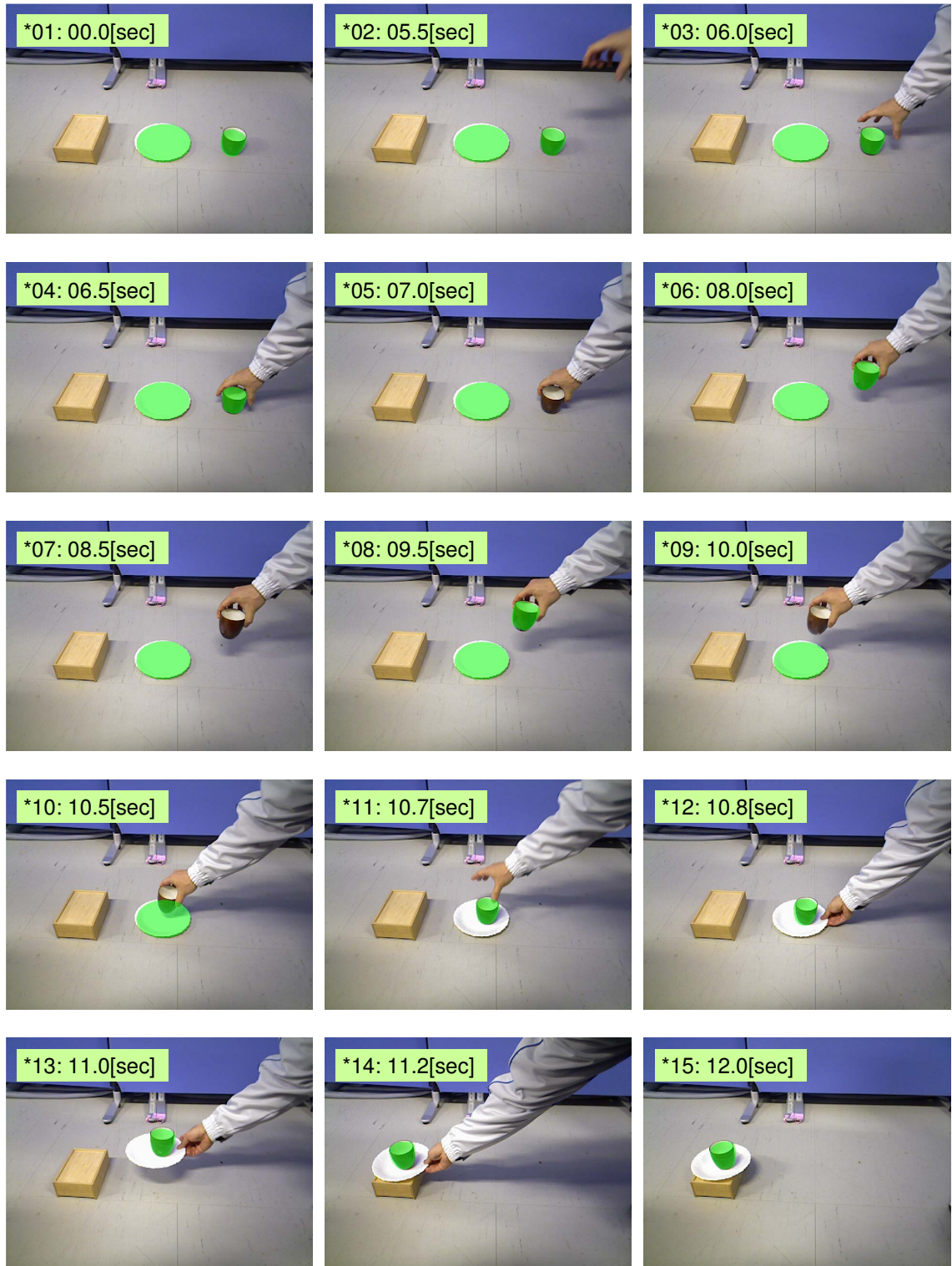


図 5.31: 予備実験 1 : LINE-MOD によるカップ・皿の追跡結果 . 対象が認識された時刻には , 緑色で対象のテンプレートを重畳表示している (著者論文 [104] にて一部抜粋した図を发表済)

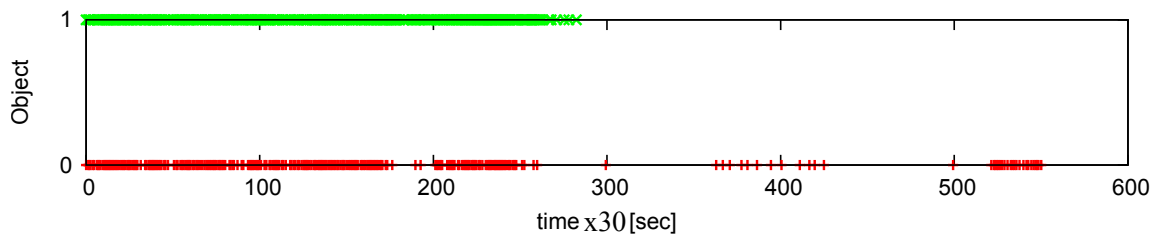


図 5.32: 予備実験 1 で LINE-MOD により対象が認識された時刻 . Object 0 がコップを , Object 1 が皿を表す (著者論文 [104] の図の時間軸の誤りを修正)

表 5.2: 改良 LINE-MOD による物体の認識特性 : 物体が上に乗っていないことを陽性と期待する場合 . Experiment 1 , 2 はそれぞれ観察 HSR-OBS1 , HSR-OBS2 での認識結果を表す

		True positive	False positive	False negative	True negative	Number of trials
Basic experiment 1	Cup	279	0	80	0	359
	Dish	317	0	1	41	359
Basic experiment 2	Cup	190	0	32	0	222
	Dish	144	0	11	67	222
Experiment 1	Cup	96	0	167	116	379
	Dish	41	0	8	330	379
Experiment 2	Cup	31	0	151	94	276
	Dish	91	0	106	79	276

察 HSR-OBS2 において , 認識対象物の遮蔽率が 20 % を超える場合を陰性 , 超えない場合を陽性と期待して偽陽性および偽陰性を計算した結果を表 5.2 に示す . いずれの結果においても , 偽陽性がゼロとなっており , 改良 LINE-MOD は , V-FES と組み合わせて利用し得る特性を持っていることがわかった .

5.6.3 ロボットによる実環境での片付け観察実験

改良 LINE-MOD と V-FES を組み合わせ , 生活支援ロボット HSR を用いた片付けタスク観察学習実験をおこなった . 実験環境を図 5.1 に示す . 実験で利用したのは , HSR の頭部 RGBD カメラ (ASUS Xtion PRO LIVE) で人が食器片付けをする

様子を撮影した画像列である．食器片付けは，以下の二通りの方法でおこなった：

観察 HSR-OBS1

机の上にコップと皿があり，人は左手でコップを，右手で皿を持って運び，いずれもサイドボードの上へ置いた．

観察 HSR-OBS2

机の上にコップと皿があり，人はまずコップを皿の上へ重ね，下の皿を持って運び，サイドボードの上へ置いた．

観察 HSR-OBS1 と HSR-OBS2 のいずれにおいても，人が机からサイドボードへ移動している際，カメラから食器がフレームアウトしている瞬間が存在している．

二種類の RGBD 画像列について，コップと皿を観察対象とし，改良 LINE-MOD を用いて追跡させた結果を図 5.33（観察 HSR-OBS1），図 5.34（観察 HSR-OBS2）に示す．

さらに，観察 HSR-OBS1 および観察 HSR-OBS2 の観察対象追跡において，観察対象の検出に成功した時刻を，図 5.35（HSR-OBS1），HSR-OBS2）に示した．図 5.35 からは，食器がフレームアウトしている時刻や，上に別の食器が重ねられている時刻で検出できていないことが分かる．また，食器が動いている時刻や，サイドボード上にあって観測点より距離が遠い時刻において，検出率が低下していることが分かる．

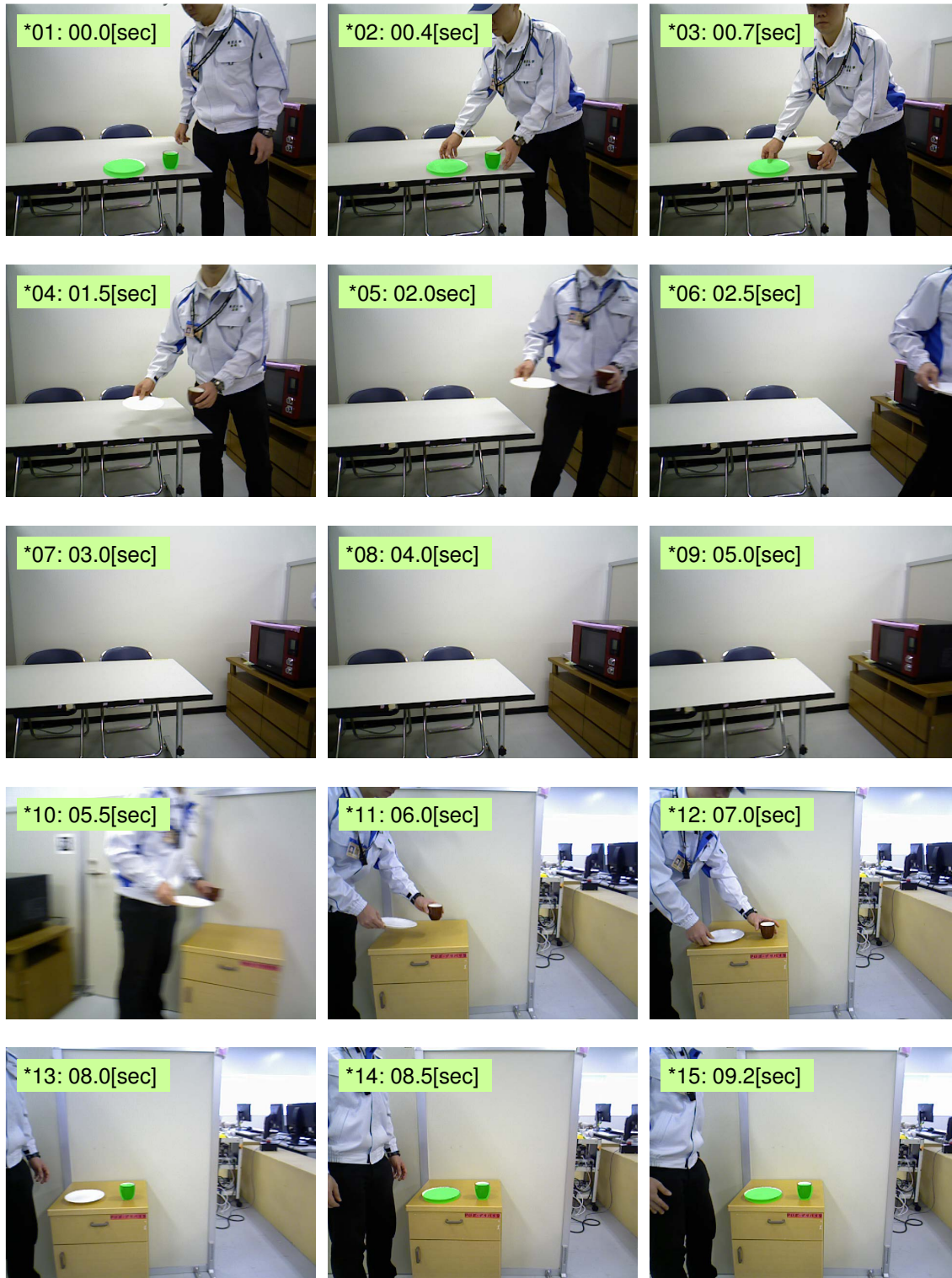
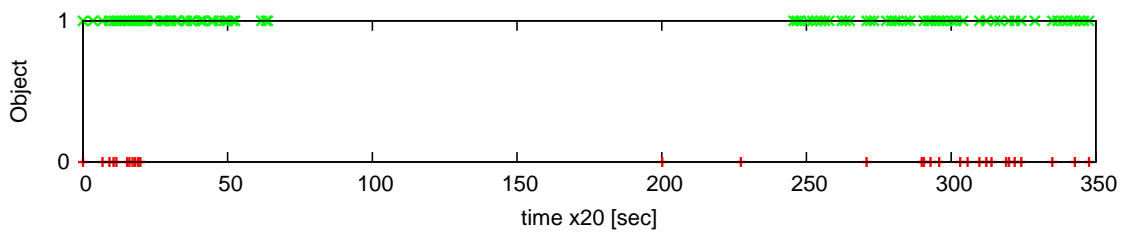


図 5.33: 改良 LINE-MOD による観察対象の追跡結果 (観察 HSR-OBS1) . 対象が認識された時刻には, 緑色で対象のテンプレートを重畳表示している (著者論文 [104] にて一部抜粋した図を发表済)



図 5.34: 改良 LINE-MOD による観察対象の追跡結果 (観察 HSR-OBS2) . 対象が認識された時刻には, 緑色で対象のテンプレートを重畳表示している (著者論文 [104] にて一部抜粋した図を发表済)

HSR-OBS1)



HSR-OBS2)

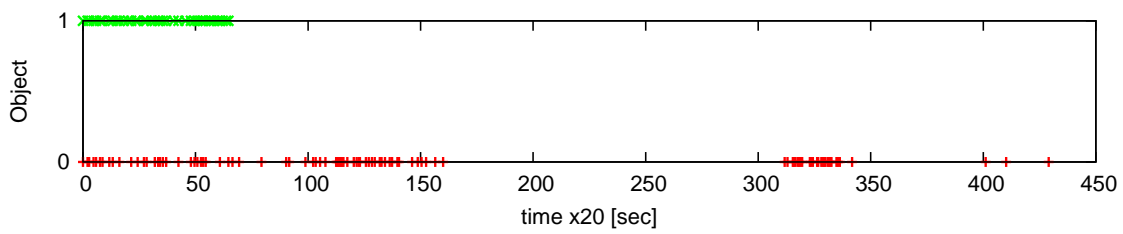


図 5.35: 観察 HSR-OBS1, HSR-OBS2 で改良 LINE-MOD により対象が認識された時刻. Object 0 がコップを, Object 1 が皿を表す (著者論文 [104] より)

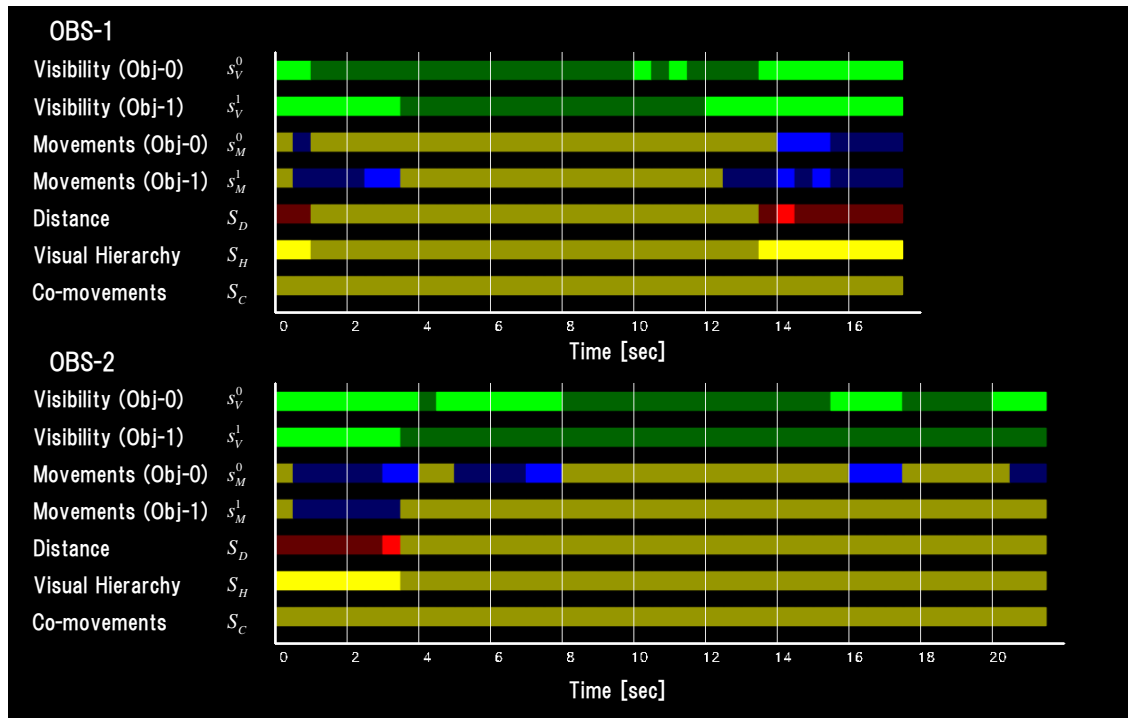


図 5.36: 観察 HSR-OBS1, 観察 HSR-OBS2 の観察対象認識結果から直接計算可能な, 物体運動, 物体同士の距離, 視覚重畳関係と運動随伴性の推定結果. 知識を利用する前の状態である (著者論文 [104] より)

5.6.4 V-FES を用いた作用・道具機能推定結果

観察対象の追跡結果を入力として, V-FES を用いて視覚重畳関係, 運動特徴を推定した過程と結果を図 5.36 から図 5.39 に示す. 評価結果とグラフに描画された色との対応は図 5.5 に示す通りとし, 図中 OBS-1 が HSR-OBS1 の結果を, OBS-2 が HSR-OBS2 の結果を表すものとする.

最終的に得られた結果 (図 5.39) によると, 観察 HSR-OBS1 では, 時刻およそ 4[sec] から 10[sec] までの, 観察対象のフレームアウトや物体の速い移動に起因する認識失敗区間においても, 二つの対象が移動している状況が正しく推定できている. 観察 HSR-OBS2 では, それに加えて, 二つの対象が途中から近づいていること, 二つの対象に片側遮蔽 (One-way) の視覚重畳関係が生じていること, さらに二つの対象に随伴する運動が見られることが正しく推定されている.

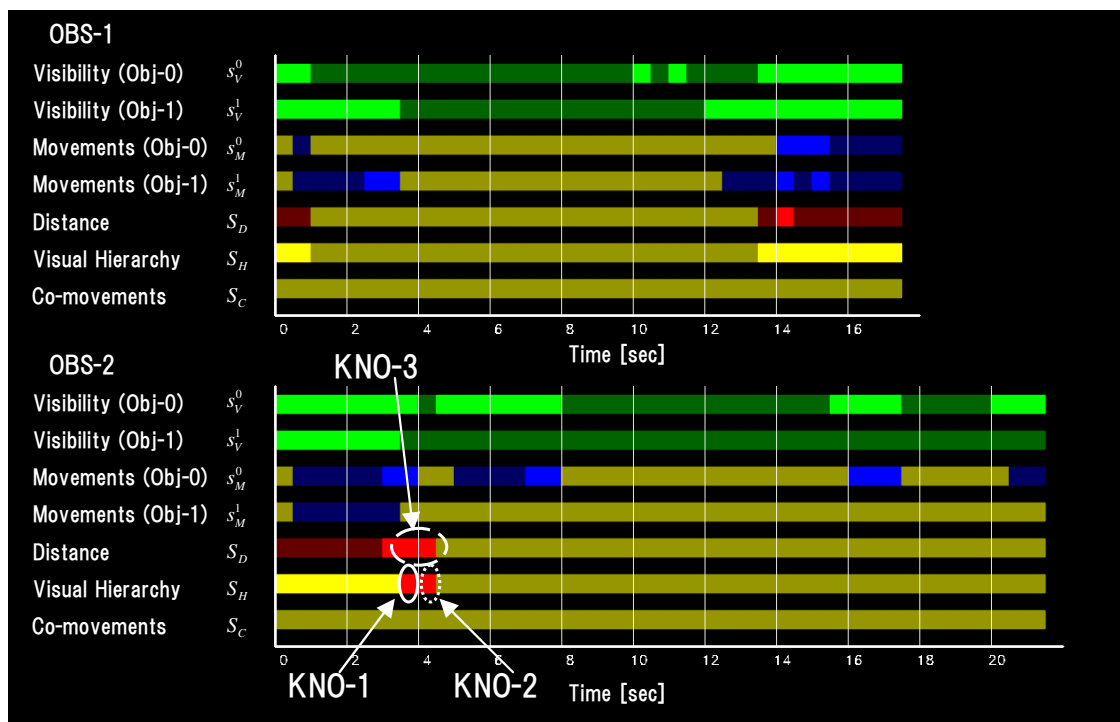


図 5.37: 観察 HSR-OBS1, HSR-OBS2 について, 知識 KNO-1 から KNO-3 までを一度だけ利用して修正した後の, メタ情報推移

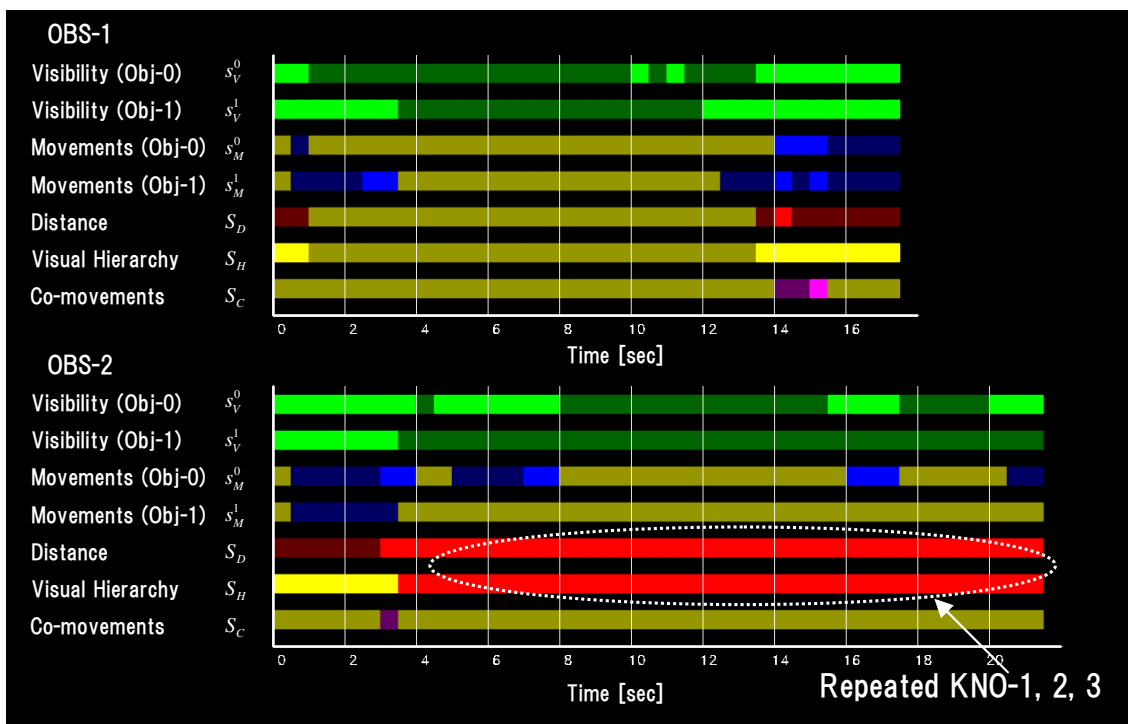


図 5.38: 観察 HSR-OBS1, HSR-OBS2 について, 知識 KNO-1 から KNO-3 までを, 収束するまで繰り返し利用して修正した後のメタ情報推移. 観察 HSR-OBS2 において, 対象同士が十分に近く, 片側遮蔽関係が存在していることが推定されている

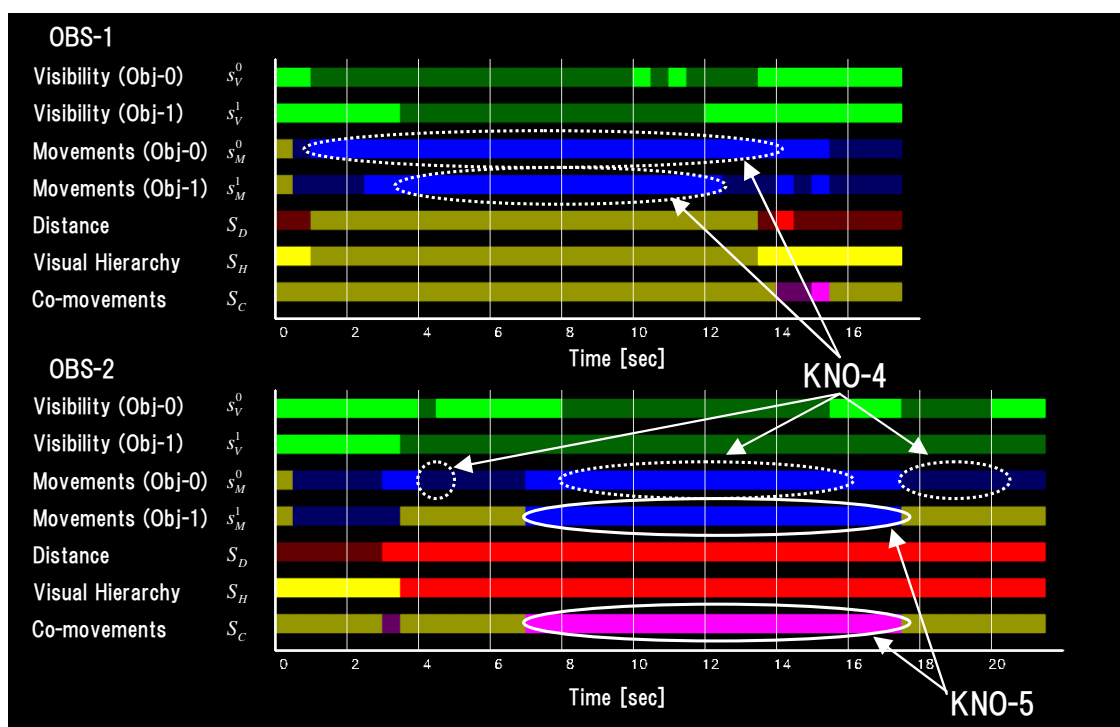


図 5.39: 観察 HSR-OBS1, HSR-OBS2 について, 運動情報を推定するとともに, 知識 KNO-4, 5 を用いてメタ情報を修正した後の, 最終的なメタ情報推移. 観察 HSR-OBS2 では二つの対象に随伴性のある運動が観測されている (著者論文 [104] より)

```
;; TASK
* OBJECT0 (CUP)
AT TIME=0.5[sec], MOVED_FROM PLACE= TABLE,
AT TIME=15.5[sec], MOVED_TO PLACE= SIDEBOARD,
* OBJECT1 (DISH)
AT TIME=2[sec], MOVED_FROM PLACE= TABLE,
AT TIME=12.5[sec], MOVED_TO PLACE= SIDEBOARD,

;; TASK GOAL:
AT OBJECT0 SIDEBOARD
AT OBJECT1 SIDEBOARD
```

図 5.40: V-FES による観察 HSR-OBS1 のタスク解釈結果

さらに本章4節で述べた手法により、観察中に表れる作用と、観察対象に表れる道具としての機能を推定した。ただし、作用推定における場所のシンボル化では、机の上およびサイドボードの上の領域を予め与えるものとした。観察 HSR-OBS1、観察 HSR-OBS2 それぞれの推定結果を図 5.40、図 5.41 に示す。観察 HSR-OBS1 では、コップと皿がそれぞれテーブルからサイドボードへ移動されたことが正しく推定されている。観察 HSR-OBS2 では、時刻 4[sec] でコップが皿に載せられ、さらに重なった状態でサイドボードへ運ばれたことが正しく推定されている。さらに観察 HSR-OBS2 では皿について、a) 上に物を載せても良い物体であること (MAT)、b) 上に物を載せた状態で運んでも良い物体であること (TRAY) が正しく推定されている。

本実験により、偽陽性のない視覚探索手法であれば、V-FES と組み合わせることにより、上下関係を変化させる操作を含むタスクを視覚学習可能であることが示された。

```
;; TASK
* OBJECT0 (CUP)
AT TIME=2.5[sec], MOVED_FROM PLACE= TABLE,
AT TIME=4[sec], LOADED_ONTO OBJECT1,
(AT TIME=6.5[sec], START being TRANSPORTED,
  by OBJECT1)
(AT TIME=17.5[sec], STOP being TRANSPORTED,
  by OBJECT1)
* OBJECT1 (DISH)
AT TIME=6.5[sec], MOVED_FROM PLACE= TABLE,
AT TIME=17.5[sec], MOVED_TO PLACE= SIDEBOARD,

;; TOOL INFORMATION
OBJECT1 CAN BE USED AS MAT
OBJECT1 CAN BE USED AS TRAY

;; TASK GOAL:
ON OBJECT0 OBJECT1
AT OBJECT1 SIDEBOARD
```

図 5.41: V-FES による観察 HSR-OBS2 のタスク解釈結果

5.7 おわりに

本章では、観察対象が完全に認識不可能となってしまう状況を含む、道具を用いたタスク観察学習について論じた。このような場合には、対象が完全に認識不可能となる時区間の前後の情報、つまり遮蔽時刻近傍情報を活用することが有効となる。複数対象の運動情報、認識の可否と視覚重置関係の時系列情報について、それらをつなぎ合わせる幾つかの知識に基づいて修正・変更をおこない、遮蔽に至る原因を推定するとともに、1) 観察対象へいつどのような作用が及ぼされているか；2) タスク目的；3) 対象の道具的機能、の三点を推定するための機構“V-FES”を提案し、その有用性を評価した。

対象が対象に載った、あるいは対象が対象に入った、そして動かされたという関係を推測するためには、

1. 「対象同士が近づいた後に一方の対象しか見えなくなったら、その後にもう一方が見えるまでは片側遮蔽関係だろう」
2. 「片側遮蔽の後に観察対象が両方見えなくなったら、手前の観察対象の認識失敗だろう」
3. 「近接した後に片側遮蔽関係が発生しているということは、物体が近いということ」
4. 「他の対象に隠されていない対象を見失ってから再度発見した場合には、単なる認識失敗である。位置が変わっていなければ「止まっている」、そうでなければ「動いている」」
5. 「物体 X が物体 Y を隠している場合、物体 X が十分動いたときにも片側遮蔽であれば、物体 Y は物体 X に随伴する運動をしてきた」

という五つの知識を用いることが重要であった。

そしてこれらの知識の適用には、それぞれ教示者や教示環境に関する仮定を必要とした。具体的には、以下のものである。

1. 対象が重要な見え隠れをする近くの時刻で、必ず対象が認識可能であること

2. タスクに重要な相互作用をしあう対象は、全てロボットが検出可能であること
3. 二つの対象と観測点が一直線上に並んだ状態で、対象同士の距離が有意に変化しないこと
4. 観察対象が消滅しないこと

逆の視点から考えると、物体を容易に見失いかねない状況での、物体の上下・包含関係を教示する際に重要な、教示者が注意すべき点が、これらの仮定の前提条件として明らかになったとも言える。これらは簡潔にまとめると「教示者は観察者へ、観察させたい上下・包含関係の変化をきちんと見せること」となるが、自然かつ重要な事項と考えられる。

以上、本手法は、完全な遮蔽下にある対象についても前後の情報を用いて重畳関係推定を可能とするものであったことから、2章で指摘した「日常的な道具と操作対象の間の遮蔽関係の観測の可能性」を肯定するものである。ただし、完全に見えない状況下での対象の細かな位置姿勢の推定はできていない。従って、軌道レベルでのタスク学習へそのまま用いることは難しいと言える。このような場合には、対象の出現・消失時刻において時刻を分節し、視覚教示結果以外の情報（触らせてもらった際の力センサ情報、あるいは音など）を用いるべきだと考える。

同じく2章で指摘した「視覚重畳関係を用いた物体同士の関係推定の可能性」についても、片付けタスク等で必須となる、物体を物体に入れる、あるいは載せる、動かす、運ぶという作業について視覚認識が可能であったことから、指摘事項を肯定する結果であったと言える。

ただし、本章の手法は、観察対象同士の上下・包含関係変化に伴って一方の対象が完全に見えなくなる場合を想定した、物体の視認性と運動情報のみを用いて片側遮蔽関係を推定するものであった。大きな布地に小さなぬいぐるみが載っている状況などは、下の物体が完全に見えなくなることは無く、この場合は前章で述べた空間遮蔽近傍情報がそのまま利用できる場合となる。このような場合分けが不要な手法とするためには、図5.29に示したように、時空間遮蔽近傍特徴を統合して用いる方法へと発展させるべきだと考えている。

V-FESが出力するタスク目的と対象の道具的機能については、STRIPSタイプのオペレータを用いたタスクプランニングに利用可能な形式で出力している。再現時

の環境やロボットの能力に応じた操作手順の変更とタスク実行へこれらの情報を活用し、各物体に許される操作のみを用いてプランニングをおこなう手法については、次章で述べる。従って、本章で提案した視覚学習法にまつわる「視覚重畳関係を用いた物体同士の関係推定を用いた動作学習の可能性」の後半および「生活支援タスクの学習・実行に際しての有用性」については次章で検証する。

第6章

作用・機能推定結果に基づくロボットの
身体に合わせたタスク再現

6.1 はじめに

前章までの認識機構により，ロボットは道具を用いたタスクの目的，そして各物体に許される道具的な利用法を，人間の動きの観察によって獲得することができた．本章では，獲得した各情報に基づき，ロボットが自身の身体と環境に合わせてタスクを再現するための方法論と，その実験結果について述べる．

一般に，タスク再現時の状況が観察学習したときと全く同じであるならば，観察と全く同じ動作をおこなえば良い．しかし通常は，動かすべき物体の位置などの状況がタスク再現状況毎に異なる．また，教示者とロボットとの体の差異によっても，動作戦略を変更しなければならない状況は多く発生する．そこで，このような場合にも，各対象に許されるような使い方のみを用いて，各状況からタスク目的を達成するための，動作の種類と順序を決定する手法が必要となる．従来よりロボットの複数の動作の手順を計算するシステムとして，また失敗した動作からの復帰をおこなえるシステムとして，作業目標レベルの高次タスクプランニングを用いる研究がおこなわれている [57, 70, 93, 124]．本研究でも同様のプランニング手法を用いるが，特に物体の道具の利用に際し，複数の物体同士の関係を記述する述語を追加するとともに，前章までの手法で推定した道具利用法，つまり物体をどのように用いてよいかという知識を確認しながら，操作手順を決定していく手法について検討，評価をおこなう．また，片付けタスクや洗濯物分類タスクにおいて本手法を用い，生活支援タスクにおける有用性について検証する．

従って本章で検証するのは，2章で述べた以下の事項である．

1. 視覚重畳関係を用いた物体同士の関係推定を用いたタスク学習の可能性
2. 生活支援タスクの学習・実行に際しての有用性

なお本章の記述は，著者論文 [104] の内容を含み，一部の実験結果については [96] でも報告している．

6.2 道具を利用したタスクの自律計画法

ロボットのタスクプランニングとしては，各状態と状態間で起こり得る遷移をシンボリックに表現し，目標状態までの遷移方法を探索する手法が用いられている．そ

して、ロボットのタスク観察学習と組み合わせてタスク再現に用いる研究は、観察学習の黎明期よりおこなわれている [44] .

Fikes らは、STRIPS (STanford Research Institute Problem Solver) と呼ばれるタスク記述言語・処理系を開発した [25] . STRIPS では、述語 (predicate) を用いて状態を表現する。状態を遷移させる演算子 (operator) としては、動作 (action) , 動作に関わるパラメタ (parameter) , 前提条件 (precondition) と副作用 (effect) のセットを予め定義しておく。ソルバは predicate を用いて表現された初期状態から目標状態へ導くような、action の列 (動作列) を探索するようになっている。

近年では、型情報 (typing) や、動作に必要なコスト (action-costs) を定義できるように、STRIPS を拡張した PDDL (Planning Domain Definition Language) 言語 [28] およびそのソルバが開発され、生活支援ロボットの複雑なタスクプランニングや失敗復帰に利用されている [57, 70, 93, 124] . 本研究でも、道具利用により生じる効率化を表現するため、PDDL 言語および、動作コストを考慮した探索が可能な PDDL のためのソルバ、 $FF(h_a)$ [42] を用いるものとした。

一つの動作単位の粒度をどのようにするかという点には課題が残るが、従来研究 [70, 93] では、各々の操作の後に失敗状態を確認し、失敗復帰が可能な粒度としており、本研究でもこの考え方を採用するものとした。

6.2.1 PDDL によるタスク記述と道具利用法表現

PDDL では、以下の二種類の記述を入力としてタスクプランニングをおこなう。

1. タスクがおこなわれる環境や operator を定義した domain ファイルと呼ばれる記述
2. タスクを始める前の初期状態や目標状態、関連する物体を定義した problem ファイルと呼ばれる記述

図 6.1 に domain ファイルの例を、図 6.2 に problem ファイルの例を示す。

domain における “:types” では、プランニングに必要な変数 (登場人物) の属性を宣言する。ロボットによる片付けプランニングでは、片付けられる物体を表す “OBJECT” , 物体が運ばれるべき場所を表す “SPOT” , ロボット自身を表す “ROBOT” ,

そしてロボットのマニピュレータを表す“ARM”の4つの属性を用いる。“:predicates”は、オペランドの状態を表す述語を定義するものである。predicatesを用いた状態表現は必ず真・偽のいずれかになる。例えば“(on? ?object - object ?object - object)”が真であるとき、“OBJECT”属性の変数二つをオペランドとして、後者の変数が指す物体の上に、もう一方の物体が載っていることが表現される。片付けプランニングで必要となる predicates については、後で述べる。“:functions”は、プランニングで必要となる関数を定義する。“:action”は、シンボリックな状態遷移を起こすための演算子 (operator) を追加するものである。operator は action 名、前提条件 (precondition)、副作用 (effect) により定義される。ただし副作用とは、その action により引き起こされる状態変化のことである。前提条件と副作用は predicates と変数を用いて表現され、表現に必要な変数とその属性は、パラメタ (parameter) として、記述される。片付けプランニングに必要な action については、後で論じる。

図6.3に示したのが、道具利用を含む片付けプランニングのための PDDL predicates である。まず“on?”は、物体の上に物体が載っていることを表すための述語である。“(on? A B)”が真のとき、物体 A が物体 B の上に載っていることを表す。ランチョンマットの上に食器が置かれた状態、あるいはトレーで皿を運んでいる状態などで真となる。物体を載せておく道具、載せて運ぶ道具を用いるプランニングで必要となる述語である。“in?”は、物体の中に物体がしまわれていることを表すための述語である。“(in? A B)”が真のとき、物体 A が物体 B の中に入っていることを表す。小物入れの中に物を収納した状態、あるいは書籍をダンボールに入れて運んでいる状態などで真となる。物体を中へ入れておく道具、入れて運ぶ道具を用いるプランニングで必要となる述語である。“at?”は、物体がある場所に存在することを表すための述語である。“(at? A S)”が真のとき、物体 A が場所 S に存在することを表す。片付けで主となる物体の移動の計画に必須であるほか、ロボットが物体を手取るために移動する計画を立てるためにも必要となる述語である。“near?”は、ロボットがある場所の近くにいて、ある場所へ手が届くことを表すための述語である。“(near? R A)”が真のとき、ロボット R がマニピュレータで、物体 A を手にとることができることを表す。ロボット自身の移動を含むタスク計画で、ロボットの位置を表すために用いる述語である。ロボット自身の移動とマニピュレータの軌道を同時に計画するならば不必要だが、この点に関しては後述する。“hold?”は、

```
01: (define (domain MANIP)
02:   (:requirements :typing :action-costs :strips :equality)
03:   (:types
04:     OBJECT SPOT ROBOT ARM)
05:   (:predicates
06:     (on? ?object - object ?object - object)
07:     ...
08:   )
09:   (:functions
10:     (length ?spot_from - spot ?spot_to - spot)
11:     (total-cost))
12:   (:action MOVE
13:     :parameters (?robot - robot ?from-spot - spot
14:                  ?to-spot - spot)
15:     :precondition (and (near? ?robot ?from-spot))
16:     :effect (and (near? ?robot ?to-spot)
17:                 (not (near? ?robot ?from-spot))
18:                 (increase (total-cost) 10))
19:   )
20:   (:action HOLD
21:     ...
22:     ...
23:   )
24: )
```

図 6.1: PDDL の domain ファイルの例


```
01: (define (problem CUPDISH)
02:   (:domain MANIP)
03:   (:objects CUP DISH - OBJECT
04:             TABLE SIDEBOARD - SPOT
05:             L-ARM - ARM
06:             HSR - ROBOT)
07:   (:init (at cup table)
08:          (at dish table)
09:          ...
10:          )
11:   (:goal (and (at cup sideboard)
12:              (at dish sideboard)
13:              ))
14:   (:metric minimize (total-cost)))
```

図 6.2: PDDL の problem ファイルの例

```
1  (:predicates
2  (on? ?object - object ?object - object)
3  (in? ?object - object ?object - object)
4  (at? ?object - object ?spot - spot)
5  (near? ?robot - robot ?spot - spot)
6  (hold? ?robot - robot ?object - object ?arm - arm)
7  (handempty? ?robot - robot ?arm - arm)
8  (clear? ?object - object)
9  (can-be-used-as-mat? ?object - object)
10 (can-be-used-as-tray? ?object - object)
11 (can-be-used-as-storage? ?object - object)
12 (can-be-used-as-container? ?object - object))
```

図 6.3: 上下・包含関係変化と道具利用を含む片付けプランニングのための PDDL predicates

ロボットがマニピュレータで、特定の物体を持っていることを表すための述語である。“(hold? R A M)” が真のとき、ロボット R がマニピュレータ M で、物体 A を持っていることを表す。物体を目的地へ置くことができる状態か、ロボットの移動と共に物体が移動する状態かを表すのに必要であり、マニピュレーション一般に必要な述語である。“handempty?” は、ロボットのマニピュレータが物を持っておらず、これから物体を持つことができる状態であることを表す述語である。“(handempty? R M)” が真のとき、ロボット R のマニピュレータ M が物を持っていないことを表す。“hold?” では、特定の物体とマニピュレータとの関係を表現していたのに対し、“handempty?” が存在することで、新たに物体を持つことのできるマニピュレータかどうかの判定のコストが大幅に削減できる利点がある。

そして、1) “can-be-used-as-mat?”、2) “can-be-used-as-tray?”、3) “can-be-used-as-storage?”、4) “can-be-used-as-container?” の4つが、物体の道具的利用法、つまり機能を表すための述語である。各々、1) 上に物を載せることができるか否か、2) 上に物を載せた状態で運ぶことができるか否か、3) 中に物を入れることができるか否か、4) 中に物を入れた状態で運ぶことができるか否か、を示すものであり、真のときにそれらの利用法が許されていることを表す。これらの述語は、前節で視覚学習した道具的利用法 “MAT”、“TRAY”、“STORAGE”、“CONTAINER” に対応するものである。

次に、これらの述語を用いて道具利用を含む片付けプランニングをおこなうための action の定義法を図 6.4 から図 6.6 に示す。“MOVE” は、ロボット自身の移動を表す action である。これはロボットが、とある現在地 (?from-spot) から目的地 (?to-spot) まで移動する動作を表す action であり、この action が終わると、副作用としてロボットの移動がおこなわれ、移動に関わるコストが増加する。従って、ロボット (?robot) とその現在地 (?from-spot)、目的地 (?to-spot) をパラメタにとる。これはロボット自身の移動とマニピュレータの動作計画を別におこなうことで生じる action である。これらを同時におこなうか別におこなうかは、タスクプランニングの粒度の問題として未解決であると考えられるが、本研究では、家事支援のタスクプランニングに関する先行研究 [57] を参考に、ロボット自身の移動はマニピュレーションと別に考えることとする。一般にロボット自身の移動は視野を大きく変えるものであり、「視野内に操作対象が収まり、認識可能か」否かをも変化さ

```

1  (:action MOVE
2      :parameters (?robot - robot ?from-spot - spot
3                  ?to-spot - spot)
4      :precondition (and (near? ?robot ?from-spot))
5      :effect (and (near? ?robot ?to-spot)
6                  (not (near? ?robot ?from-spot))
7                  (increase (total-cost)
8                          (length ?from-spot ?to-spot)))
9  )
10 (:action HOLD
11     :parameters (?robot - robot ?object - object
12                 ?arm - arm ?spot - spot)
13     :precondition (and (near? ?robot ?spot)
14                       (handempty? ?robot ?arm)
15                       (at? ?object ?spot)
16                       (or (clear? ?object)
17                           (and (not (clear? ?object))
18                               (or
19                                   (can-be-used-as-tray? ?object)
20                                   (can-be-used-as-container? ?object))))))
21     :effect (and (increase (total-cost) 1)
22                 (not (handempty? ?robot ?arm))
23                 (hold? ?robot ?object ?arm)
24                 (not (at? ?object ?spot)))
25 )
26 (:action PLACE
27     :parameters (?robot - robot ?object - object
28                 ?arm - arm ?spot - spot)
29     :precondition (and (hold? ?robot ?object ?arm)
30                       (near? ?robot ?spot))
31     :effect (and (handempty? ?robot ?arm)
32                 (not (hold? ?robot ?object ?arm))
33                 (at? ?object ?spot)
34                 (increase (total-cost) 1))
35 )

```

図 6.4: 道具利用を含む片付けプランニングのための PDDL action 定義 (1)

```

36  (:action HOLD-ON-OBJ
37      :parameters (?robot - robot ?object - object
38                  ?arm - arm ?object_from - object
39                  ?spot - spot)
40      :precondition (and (at? ?object_from ?spot)
41                        (near? ?robot ?spot)
42                        (handempty? ?robot ?arm)
43                        (on? ?object ?object_from))
44      :effect (and (not (handempty? ?robot ?arm))
45                 (hold? ?robot ?object ?arm)
46                 (not (on? ?object ?object_from))
47                 (clear? ?object_from)
48                 (increase (total-cost) 1))
49  )
50  (:action PUT-OUT-OF-OBJ
51      :parameters (?robot - robot ?object - object
52                  ?arm - arm ?object_from - object
53                  ?spot - spot)
54      :precondition (and (at? ?object_from ?spot)
55                        (near? ?robot ?spot)
56                        (handempty? ?robot ?arm)
57                        (in? ?object ?object_from))
58      :effect (and (not (handempty? ?robot ?arm))
59                 (hold? ?robot ?object ?arm)
60                 (not (in? ?object ?object_from))
61                 (clear? ?object_from)
62                 (increase (total-cost) 1))
63  )

```

図 6.5: 道具利用を含む片付けプランニングのための PDDL action 定義 (2)

```
64 (:action PLACE-ON-OBJ
65     :parameters (?robot - robot ?object - object
66                 ?arm - arm ?object_onto - object
67                 ?spot - spot)
68     :precondition (and (hold? ?robot ?object ?arm)
69                       (at? ?object_onto ?spot)
70                       (near? ?robot ?spot)
71                       (can-be-used-as-mat?
72                         ?object_onto)
73                       (clear? ?object_onto))
74     :effect (and (handempty? ?robot ?arm)
75                (not (hold? ?robot ?object ?arm))
76                (on? ?object ?object_onto)
77                (increase (total-cost) 1)
78                (not (clear? ?object_onto)))
79 )
80 (:action PUT-INTO-OBJ
81     :parameters (?robot - robot ?object - object
82                 ?arm - arm ?object_into - object
83                 ?spot - spot)
84     :precondition (and (hold? ?robot ?object ?arm)
85                       (at? ?object_into ?spot)
86                       (near? ?robot ?spot)
87                       (can-be-used-as-storage?
88                         ?object_into)
89                       (clear? ?object_into))
90     :effect (and (handempty? ?robot ?arm)
91                (not (hold? ?robot ?object ?arm))
92                (in? ?object ?object_into)
93                (increase (total-cost) 1)
94                (not (clear? ?object_into)))
95 )
```

図 6.6: 道具利用を含む片付けプランニングのための PDDL action 定義 (3)

せるものである。操作対象の存在可否によって、動作の有無が決定されるべきことから、マニピュレーションと移動は別に扱い、その間に環境認識過程を挟むべきだと考える。ただしこの場合、“MOVE”の後のマニピュレーションに成功するような場所への移動が保証されていないといけない。しかしこの問題は、マニピュレーションに失敗すれば成功するまでロボット自身の位置を修正しながら失敗復帰をする枠組みを追加実装することで解決できると考えられる。“MOVE”を実際に行うためには、ロボットの脚部あるいは台車の動作計画が別途必要となる。

“HOLD”と“PLACE”は、それぞれロボットが手を伸ばして対象をつかみ上げる、ある場所へ置くための action である。いずれもロボットがマニピュレーション可能な場所、つまり物が置いてある場所あるいは物を置くべき場所の近くにいることが必要条件となるほか、“HOLD”の前には手が塞がっていないこと、“PLACE”の前にはきちんと目的の物をつかんでいることも precondition として確認する。“HOLD”と“PLACE”の実行のためには、ロボットのマニピュレーション動作計画、つまり逆運動学演算が必要となる。

“HOLD-ON-OBJ”、“PUT-OUT-OF-OBJ”、“PLACE-ON-OBJ”、そして“PUT-INTO-OBJ”は、ロボットがある物体の上に乗っている物体をつかみ上げる（下ろす）、ある物体の中に入っている物体をつかみ上げる（出す）、ある物体の上に物体を置く（載せる）、そしてある物体の中に物体を置く（入れる）という動作を表す action である。これらの action と“HOLD”あるいは“PLACE”の差異は、物体を操作するための目標位置が、対となって利用される物体の形状あるいは位置によって決定されるだけでなく、載せられる・入れられる対象に制限が存在することである。従って、action は載せられる物体や入れられる物体をも parameter の一つとして必要となり、action を適用するための前提条件として、それぞれ“can-be-used-as-mat?”、“can-be-used-as-tray?”、“can-be-used-as-storage?”、あるいは“can-be-used-as-container?”が真である物体であることを確認する必要がある。“PLACE-ON-OBJ”、“PUT-INTO-OBJ”ではその物体の位置が逆運動学演算に利用され、それ以外の action でも、物体同士の関係（“on”あるいは“in”）を更新し記憶し直すために必要となる。これらの action にも、それぞれ同様にマニピュレーションの動作計画として逆運動学演算が必要である。

6.3 道具を用いたタスクの再現シミュレーション実験

本節では、観察学習結果を用いたタスクプランニング手法が、ロボットの身体に合ったタスク計画をおこなうことが可能か、そして対象を観察した際に推定した、許される道具としての使い方のみを使い、効率的なタスク遂行が可能となっているかの検証をおこなう。前者に関しては、観察学習が身体の変化に頑強で再利用性があるものとなっているかを評価している事にもなる。前節までで述べたタスクプランニング手法を用い、前章でおこなった、改良 LINE-MOD と V-FES を用いた食器片付け観察学習（観察 HSR-OBS1 および観察 HSR-OBS2）の実験結果から、シミュレーション環境でロボットにタスクを再現させる実験をおこなうものとした。さらに、複数回の観察学習に基づくタスク目的の変更法についても検討する。

実験で用いるロボットは、“HSR” と “HRP-2VZ” の二種類である。HRP-2VZ に関しては3章で記述したとおりだが、HSR は作動二輪方式の移動台車と7軸アームを有し、直動伸縮式昇降機構により、アームを昇降することが可能なロボットである（[92, 115], 図 5.1）。複数物体の操作においては、HSR と HRP-2VZ では、腕の数が異なることが最大の差異だと言える。

シミュレーション環境では、ロボットはピックアンドプレースをおこなうために必要な “HOLD”, “PLACE”, あるいは物体の上に物体を置く “PLACE-ON-OBJ” 等、そして HSR 自身が作業場所へ移動するための “MOVE” に対応する動作を知識として有している。それぞれの動作では、コップや皿に適したマニピュレーション方法が選択されるものとし、各動作で失敗は無いものとする。また、マニピュレーションに必要な対象の三次元位置の特定、家具位置の特定（ローカライゼーション）には全て成功しているものとする。

6.3.1 単腕ロボット HSR でのタスク再現実験

まずはじめに、単腕ロボット HSR を用いて、観察学習したタスクを再現するシミュレーション実験をおこなった。実験開始の状況として、初期状態で HSR はサイドボード近くにおり、机の上には皿とコップが独立に置かれているものとする。さらに、HSR は初期状態の皿とコップを認識できているものとする。ただし、前述の

```

0: (MOVE HSR SIDEBOARD TABLE)
1: (HOLD HSR CUP L-ARM TABLE)
2: (MOVE HSR TABLE SIDEBOARD)
3: (PLACE HSR CUP L-ARM SIDEBOARD)
4: (MOVE HSR SIDEBOARD TABLE)
5: (HOLD HSR DISH L-ARM TABLE)
6: (MOVE HSR TABLE SIDEBOARD)
7: (PLACE HSR DISH L-ARM SIDEBOARD)

```

```
Total cost of plan: 44.000000
```

図 6.7: HSR のシミュレーション実験 1：観察 HSR-OBS1（皿とコップを独立に運ぶのを観察した場合）に基づく HSR のタスクプランニング結果

action に要するコストについて、ロボット自身が動く“MOVE”には、物体のマニピュレーションよりも大きなコストがかかるよう定義してある。

HSR のシミュレーション実験 1 まず、前章の観察 HSR-OBS1 の結果に基づく、HSR のタスクプランニング結果を図 6.7 に示す。人間は食器を両腕で運んでいたが、HSR は単腕であるため、机とサイドボードを往復しながら食器を運ぶ動作が計算された。このタスクプランニング結果に基づき、対象状態レベルと身体制御レベルの動作計画をおこないながら実行させた結果を図 6.10 に示す。

HSR のシミュレーション実験 2 次に、前章の観察 HSR-OBS2 の結果に基づく、HSR のタスクプランニング結果を図 6.8 に示す。観察 HSR-OBS2 では、人間はコップを皿に重ねてサイドボードへ運んでおり、ロボットも皿の道具的利用法“MAT”と“TRAY”を観察により獲得していた。結果、まずコップを皿の上へ置き、皿を持ってサイドボードへ運ぶという動作が計画された。タスク実行結果を図 6.11 に示す。

HSR のシミュレーション実験 3 ここで、観察学習の変更法についても検討する。HSR-OBS1 によるタスク目的の観察学習結果は (AT CUP SIDEBOARD) かつ (AT


```
0: (MOVE HSR SIDEBOARD TABLE)
1: (HOLD HSR CUP L-ARM TABLE)
2: (PLACE-ON-OBJ HSR CUP L-ARM DISH TABLE)
3: (HOLD HSR DISH L-ARM TABLE)
4: (MOVE HSR TABLE SIDEBOARD)
5: (PLACE HSR DISH L-ARM SIDEBOARD)
```

```
Total cost of plan: 22.000000
```

図 6.8: HSR のシミュレーション実験 2：観察 HSR-OBS2 (コップを皿に重ねて運ぶのを観察した場合) に基づく HSR のタスクプランニング結果

DISH SIDEBOARD) であった。一方、HSR-OBS2 によるタスク目的の観察学習結果は (ON CUP DISH) かつ (AT DISH SIDEBOARD) であり、同時に皿の道具的利用法として “can-be-used-as-mat” と “can-be-used-as-tray” が推定されている。このように、道具的利用法とタスク目的は独立に記述できていることから、タスク目的が変更されても、過去に見たことのある道具的利用法を用いて効率的なタスクの遂行が可能であると考えられる。

ここで、HSR-OBS2 の後に HSR-OBS1 を観察学習した場合の観察学習結果として観察 HSR-OBS3 を作成することを考えると、タスク目的を上書きして (AT CUP SIDEBOARD) かつ (AT DISH SIDEBOARD) とし、道具的利用法 “can-be-used-as-mat” と “can-be-used-as-tray” は引き続き用いるなどということが可能である。

このようにした観察 HSR-OBS3 に基づくタスクプランニングをおこなった。タスクプランニング結果、実行結果をそれぞれ図 6.9 と図 6.12 に示す。このプランニングでは、最初に学習した皿の道具的利用法を用いることが選択され、ロボットはコップを皿に重ねてサイドボードまで運んだが、その後コップをサイドボードの上に置き直し、後に学習したタスク目的を達成した。

以上の実験により、以下の事項が確認できた。

- タスク目的・道具的利用法の観察学習結果の利用について
皿とコップを運び、そして重ねるといったタスク目的が、シミュレーションではあるものの、実験 1 から実験 3 まで、各々達成できている。そしてその過程では、各物

```

0: (MOVE HSR SIDEBOARD TABLE)
1: (HOLD HSR CUP L-ARM TABLE)
2: (PLACE-ON-OBJ HSR CUP L-ARM DISH TABLE)
3: (HOLD HSR DISH L-ARM TABLE)
4: (MOVE HSR TABLE SIDEBOARD)
5: (PLACE HSR DISH L-ARM SIDEBOARD)
6: (HOLD-ON-OBJ HSR CUP L-ARM DISH SIDEBOARD)
7: (PLACE HSR CUP L-ARM SIDEBOARD)

```

```
Total cost of plan: 24.000000
```

図 6.9: HSR のシミュレーション実験 3：観察 HSR-OBS1 と観察 HSR-OBS2（カップと皿を重ねて運ぶのを観察した後に、皿とカップを独立に運んでいるのを観察した場合）に基づく HSR のタスクプランニング結果

体に関して学習時に観察した道具的利用法，つまり利用して良い使い方の範囲内でおこなわれていることが，実験 1 と実験 2 を比較することで実証されている．また，タスク目的と道具的利用法を別々に扱うことで，実験 3 のような，過去の道具利用法を用いて，新たに与えられたタスクをおこなうことも可能となっている．

- 体の違いに関する観察学習結果の再利用性について

作業目標レベルでの因果関係記述とタスク目的記述を利用していることから，実験 1 では，教示している人間とは腕の本数が異なる HSR においても，操作手順を変更することで，目的となるタスクを達成できている．実験 2 では，観察学習した道具的利用法を用いることで，机とサイドボードを往復しながら一つずつ物を運ぶよりも手順の少ないタスク計画がなされ，道具利用をおこなうことでの効率性向上が見られている．

6.3.2 双腕ロボット HRP-2VZ でのタスク再現実験

次に，双腕ロボット HRP-2VZ を用いたタスク再現実験をおこなった．こちらも HSR のシミュレーション実験 1 から 3 と同様に，初期状態で HRP-2VZ はサイドボード近くにおり，机の上には皿とカップが独立に置かれているものとする．

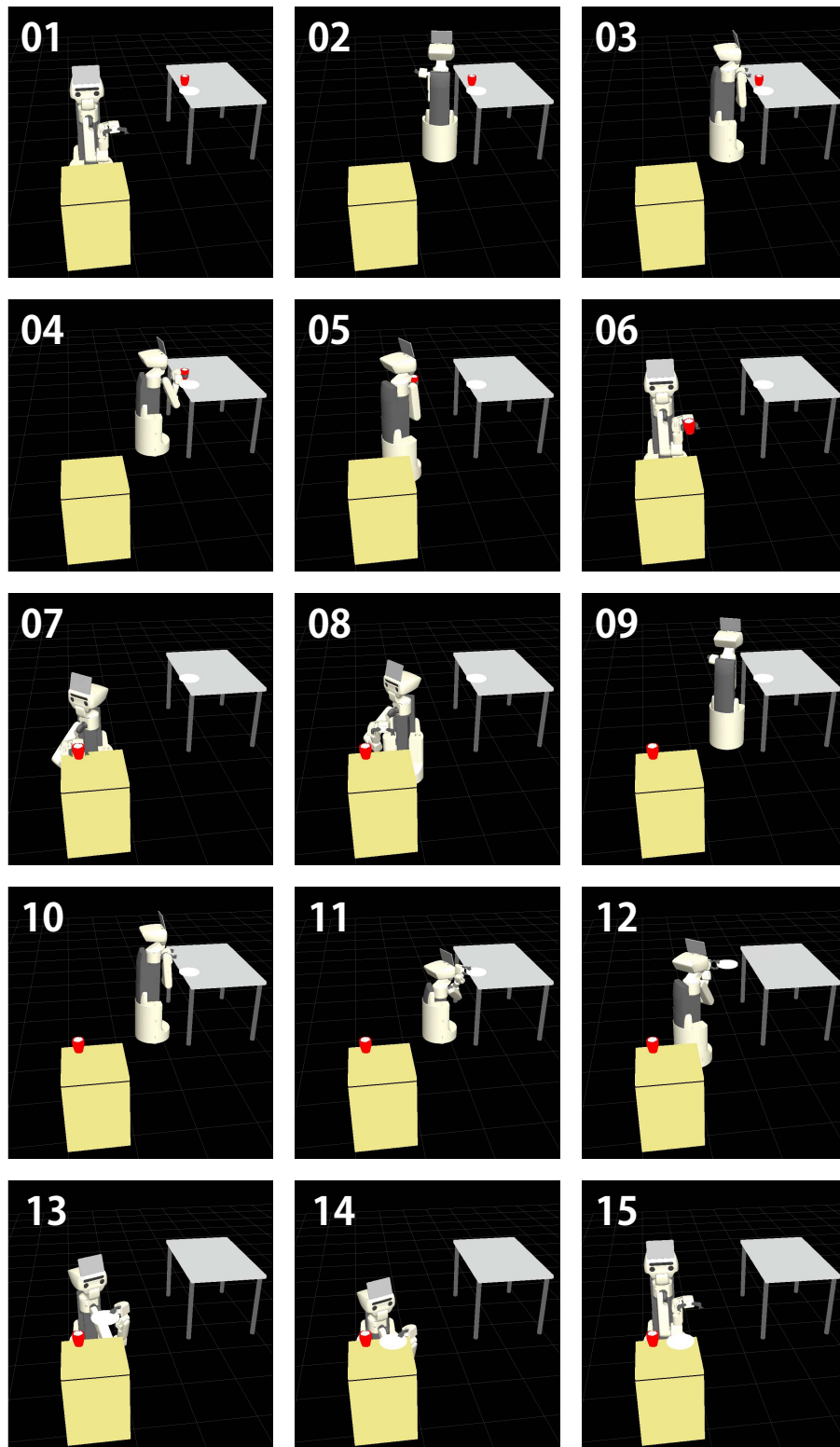


図 6.10: HSR のシミュレーション実験 1 : コップと皿を独立に運んだ観察学習結果に基づく再現結果 (一部抜粋した図は著者文献 [104] で発表済)

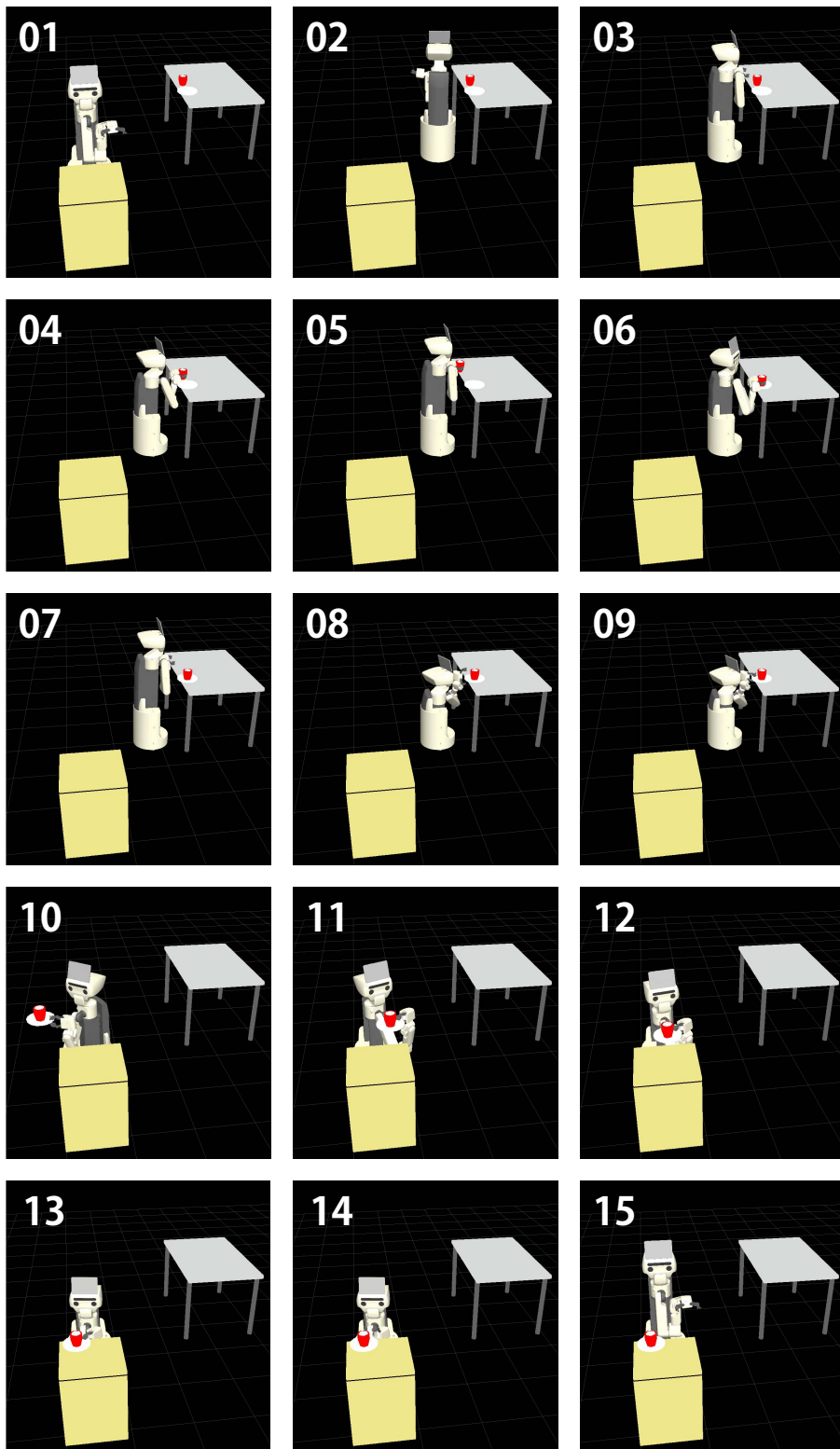


図 6.11: HSR のシミュレーション実験 2 : コップを皿に重ねて運んだ観察学習結果に基づく再現結果 (一部抜粋した図は著者文献 [104] で発表済)

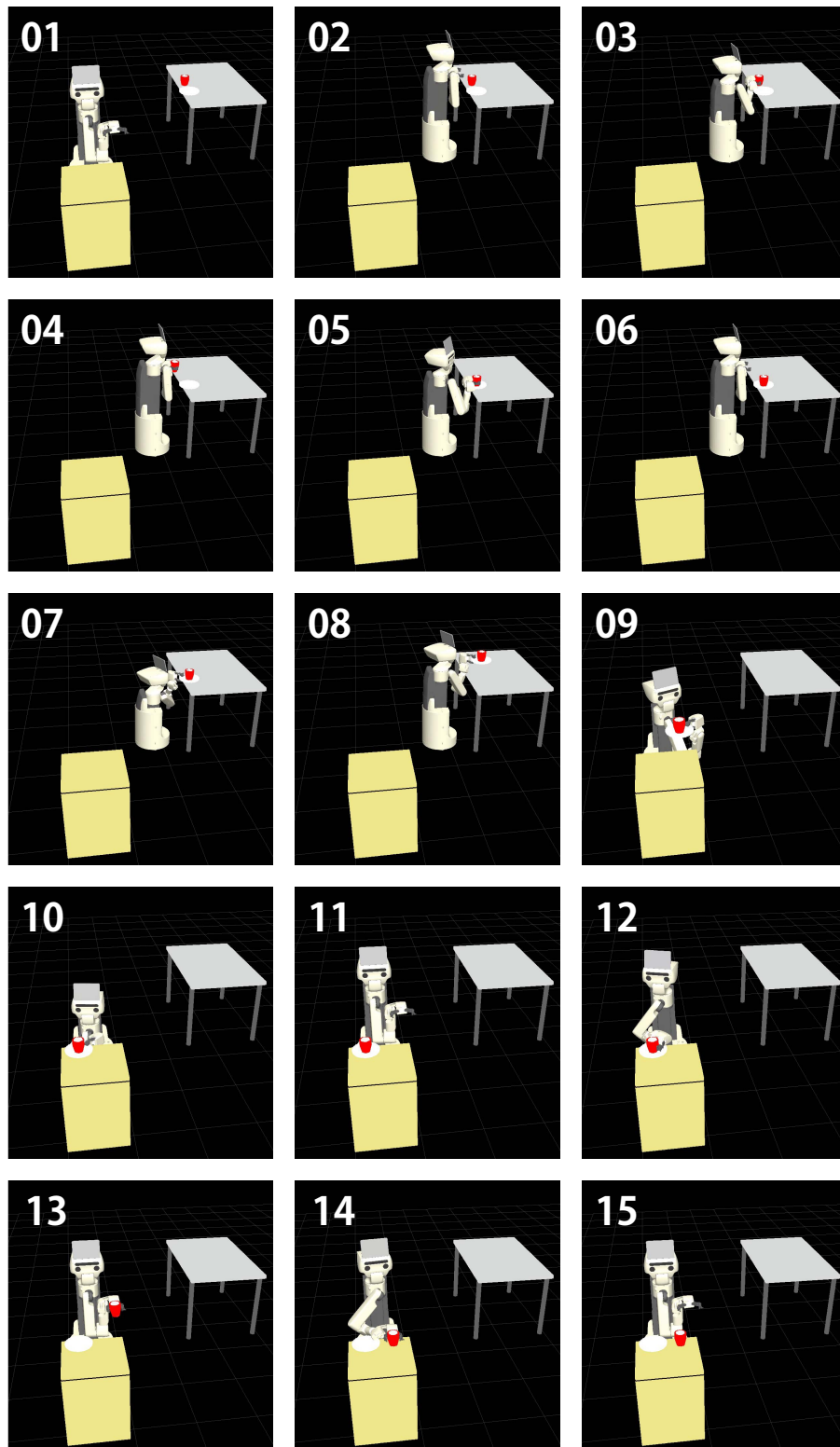


図 6.12: HSR のシミュレーション実験 3 : コップを皿に重ねて運ぶのを観察した後
に, 皿とコップを独立に運んでいるのを観察学習した結果に基づく再現結果

```
0: (MOVE HRP-2 SIDEBOARD TABLE)
1: (HOLD HRP-2 CUP R-ARM TABLE)
2: (HOLD HRP-2 DISH L-ARM TABLE)
3: (MOVE HRP-2 TABLE SIDEBOARD)
4: (PLACE HRP-2 CUP R-ARM SIDEBOARD)
5: (PLACE HRP-2 DISH L-ARM SIDEBOARD)

Total cost of plan: 24.000000
```

図 6.13: HRP-2VZ のシミュレーション実験 1 (観察 HSR-OBS1 に基づく HRP-2VZ のタスク再現) のタスクプランニング結果

```
0: (MOVE HRP-2 SIDEBOARD TABLE)
1: (HOLD HRP-2 CUP R-ARM TABLE)
2: (PLACE-ON-OBJ HRP-2 CUP R-ARM DISH TABLE)
3: (HOLD HRP-2 DISH R-ARM TABLE)
4: (MOVE HRP-2 TABLE SIDEBOARD)
5: (PLACE HRP-2 DISH R-ARM SIDEBOARD)

Total cost of plan: 22.000000
```

図 6.14: HRP-2VZ のシミュレーション実験 2 (観察 HSR-OBS2 に基づく HRP-2VZ のタスク再現) のタスクプランニング結果

HRP-2VZ のシミュレーション実験 1 観察 OBS-1 の結果に基づいて HRP-2VZ のタスクプランニングをおこなった結果，図 6.13 のようになった．タスク実行結果を図 6.16 に示す．HRP-2VZ は双腕であるため，HSR での実験とは異なり，両手でコップと皿を持って運ぶ，というプランニング結果となった．

HRP-2VZ のシミュレーション実験 2 次に，観察 HSR-OBS2 に基づくタスクプランニング結果を図 6.14 に示す．ロボットが双腕であることから，コップと皿を独立に運ぶプランが計算されても良いが，この計算ではコップを皿に載せてから運ぶというプランが計算された．タスク実行結果を図 6.17 に示す．

```
0: (MOVE HRP-2 SIDEBOARD TABLE)
1: (HOLD HRP-2 CUP L-ARM TABLE)
2: (HOLD HRP-2 DISH R-ARM TABLE)
3: (MOVE HRP-2 TABLE SIDEBOARD)
4: (PLACE HRP-2 CUP L-ARM SIDEBOARD)
5: (PLACE HRP-2 DISH R-ARM SIDEBOARD)
```

```
Total cost of plan: 22.000000
```

図 6.15: HRP-2VZ のシミュレーション実験 3 (観察 HSR-OBS3 に基づく HRP-2VZ のタスク再現) のタスクプランニング結果

HRP-2VZ のシミュレーション実験 3 最後に、コップと皿を重ねて運ぶのを観察した後に、皿とコップを独立に運んでいるのを観察させた場合 (HSR-OBS3) のタスクプランニング結果を、図 6.15 に示す。タスク目的としてサイドボード上にコップと皿が独立に置かれていることが設定され、その過程で皿をトレーのように道具として用いて良いことが学習されているが、HRP-2VZ は双腕であり、皿の道具的利用によってむしろタスクのコストが増大することから、はじめから皿とコップを独立に運んでサイドボードへ置くというプランが計算された。実行結果を図 6.18 に示す。

以上の結果は、観察学習時のロボットとタスク実行時のロボットの身体が異なっても、それぞれの身体において可能な動作と、その因果関係記述を用いれば、同様にタスク遂行が可能であるという、観察学習についての再利用性を示している。

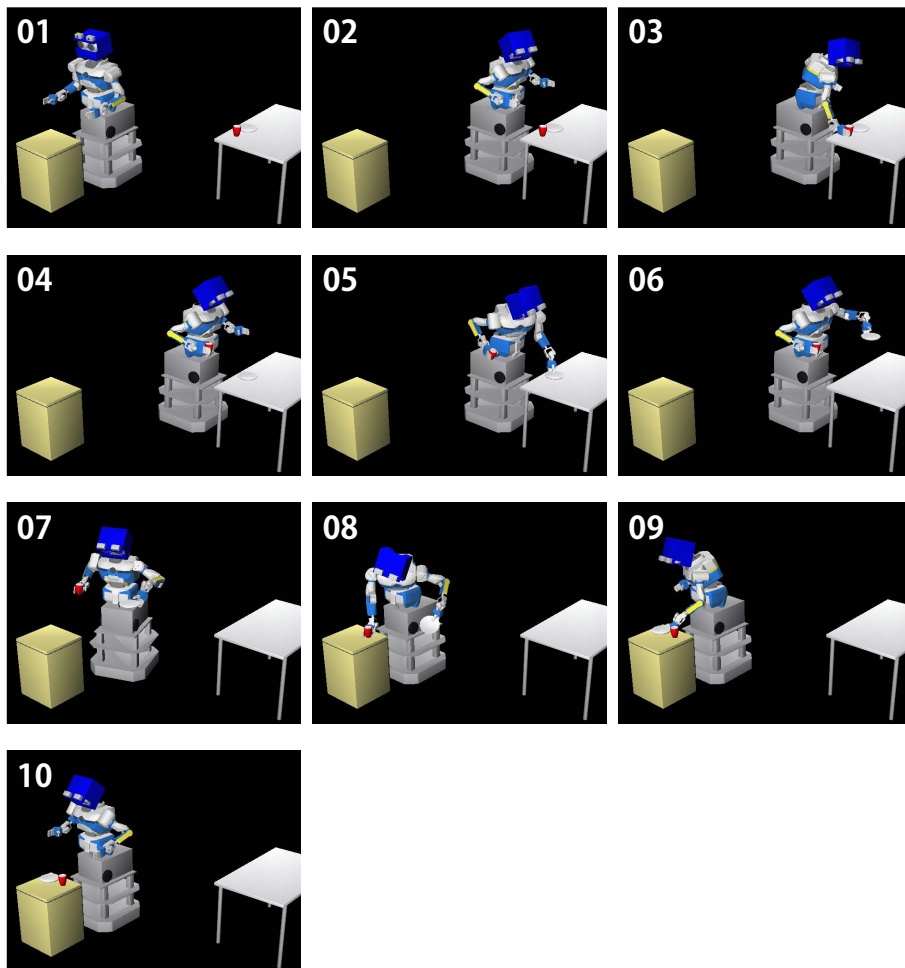


図 6.16: HRP-2VZ のシミュレーション実験 1 (観察 HSR-OBS1 に基づく HRP-2 のタスク再現) の結果

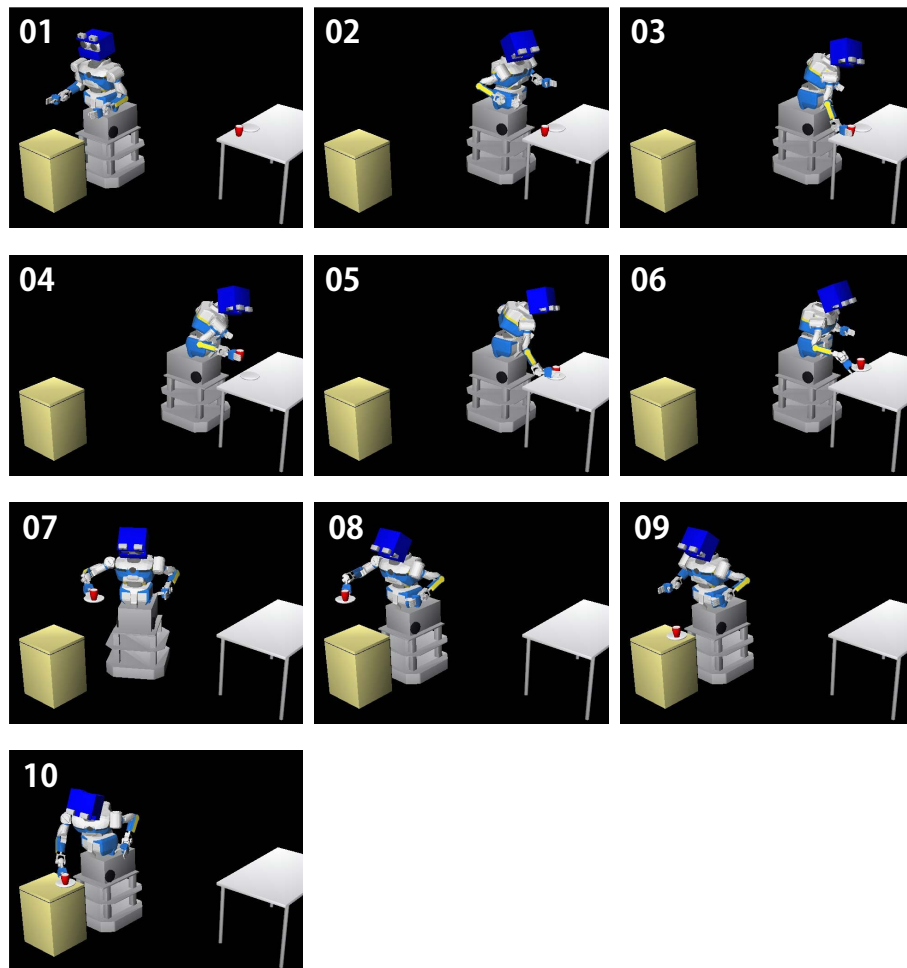


図 6.17: HRP-2VZ のシミュレーション実験 2 (観察 HSR-OBS2 に基づく HRP-2 のタスク再現) の結果

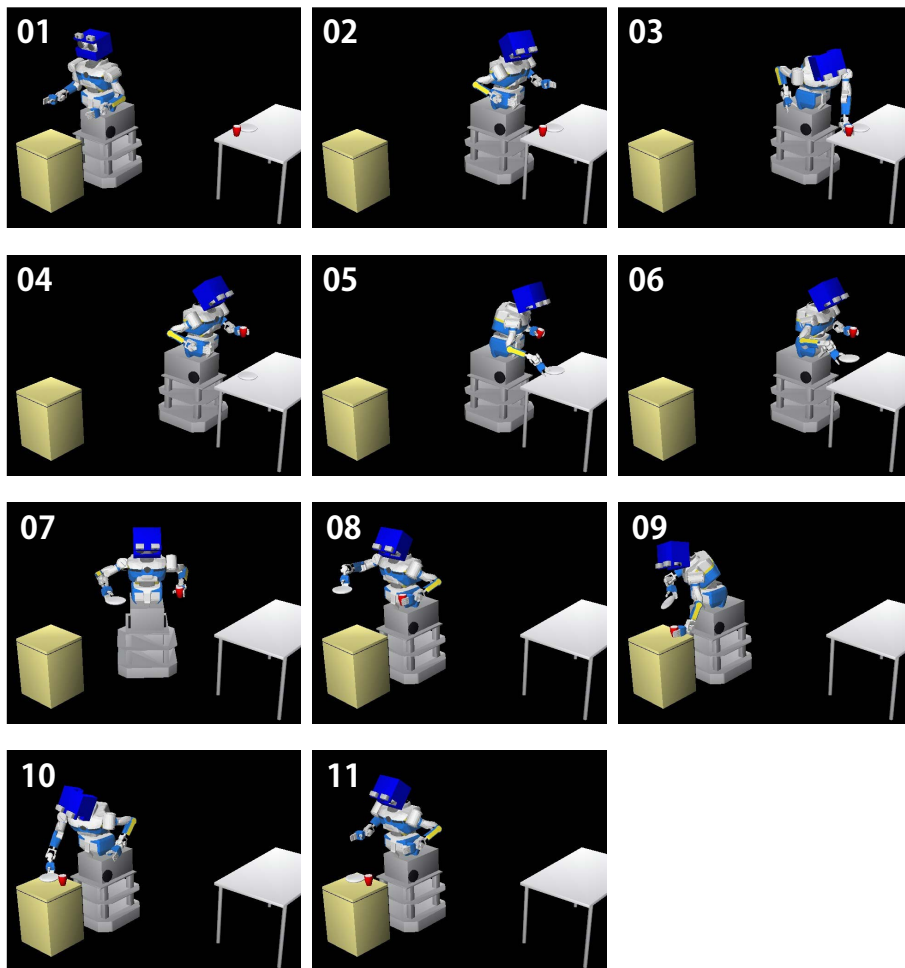


図 6.18: HRP-2VZ のシミュレーション実験 3 (観察 HSR-OBS3 に基づく HRP-2 のタスク再現) の結果

6.4 視覚重畳関係を用いる観察学習法の日常タスクにおける利用

本節では、本研究で提案する遮蔽近傍情報を用いた作用推定に基づく道具利用タスク学習法の有用性を、実タスクの実現性から評価するため、片付けシーンおよび洗濯シーンの二つを取り上げる。これらのシーンでは、従来の学習法では難しかった物体の包含関係の変化と、それに起因する見え隠れが生じる。前半では、物品を箱（引き出し）に収納するタスクを HRP-2 に、後半では洗濯物を分類して洗濯ネットを活用するタスクを HIRO に学習させ、再現させる実験をおこなう。複数回の観察学習結果の統合法、そして再現時の対象の有無と個数によるタスク目的の変更法についても検討し評価する。

6.4.1 収納に関する観察学習とその再現

まず HRP-2VZ (Mk.II) を用いて、収納・片付けシーンでの観察学習実験をおこなった。自律ロボットによる片付けとしては、部屋の中の衣類の片付けや家具の整列 [79]、皿の片付け [132]、道具の片付け [147] などが研究されており、研究例ではその物体探索方法に主眼が置かれている。しかし一般に片付けタスクでは、物体が重なり合う際、あるいは物体が物体の中へ入る際、下や中の対象が見えなくなることが多く、対象が見えない最中に何がおこなわれているかを推定するための視覚機能が必要だと言える。

観察学習時に外部カメラから撮影した画像を図 6.19 に示す。観察は二回（観察 HRP2-OBS3, HRP2-OBS4）おこなうものとし、観察 HRP2-OBS3 では机の上のタオルを箱へ入れてから棚へしまうものとした。観察 HRP2-OBS4 ではタオルは机の上に残したまま、棚へ箱を運ぶ。いずれの観察においても、タオルは観察中にロボットから認識不可能となるため、箱に入ったのか、フレームアウトなのか等の理由を適切に推定することが必要となる。観察対象をタオル (OBJECT0) と箱 (OBJECT1) とし、二回の学習結果に応じたロボットの動作再現結果を評価した。

柔軟物であるタオルの認識には、前章の観察 HRP2-OBS1/2 と同じ手法を用いるものとした。また、箱にも二種類のカラーマーカを添付し、同様の認識手法により

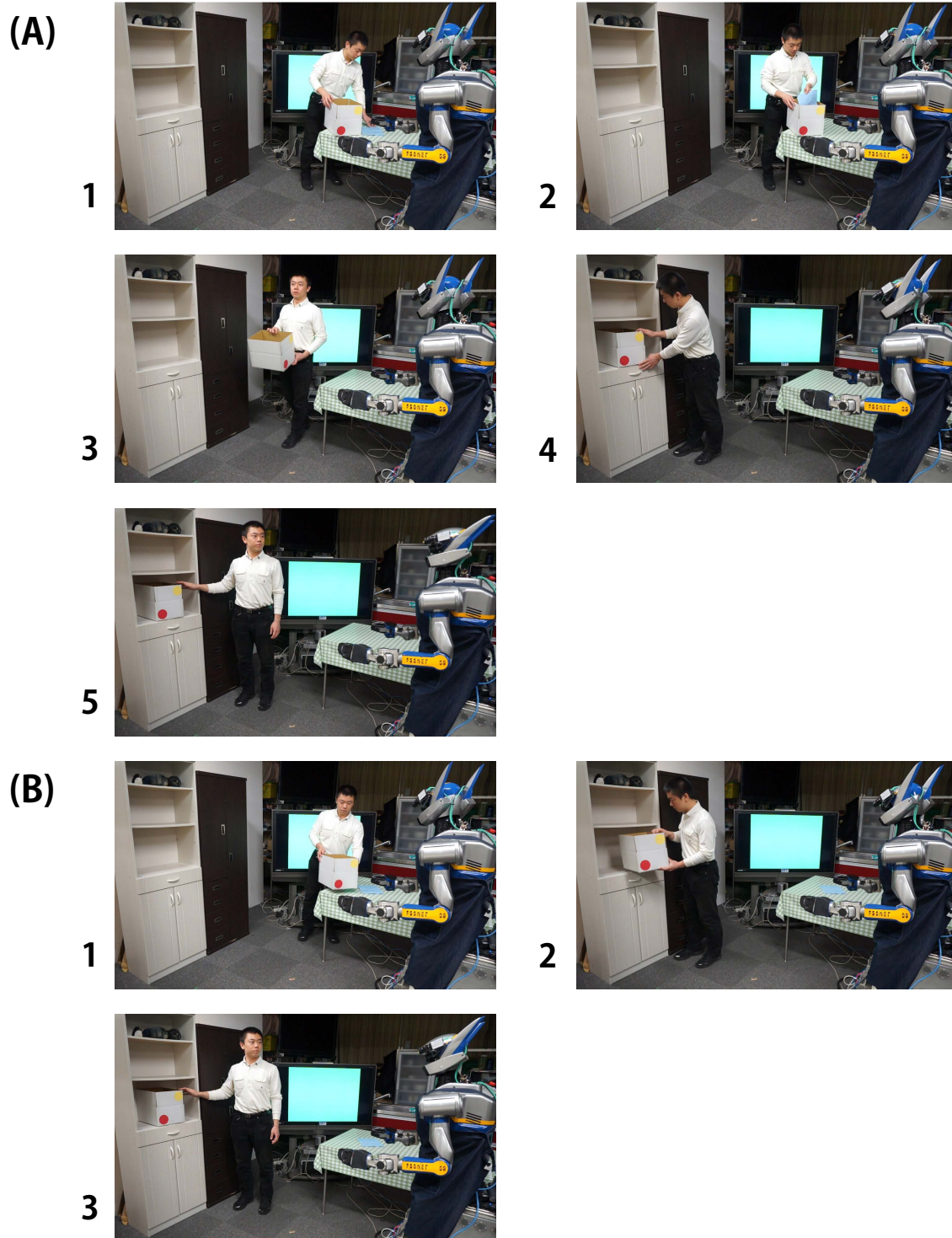


図 6.19: 収納に関する観察学習環境．(A) は観察 HRP2-OBS3，(B) は観察 HRP2-OBS4 の様子を外部カメラで撮影した様子である

位置を特定させるものとした。引き出しのような、家庭環境内への出入りの少ない物品では、QRコード等を用いて認識し利用する研究も存在する [115]。この手法も本節の認識手法と同様に、V-FES で前提とする偽陽性の少ない対象認識結果を出力するものであることから、本研究で提案する手法と併用することで、本実験結果と同等の結果が得られると考えられる。

このようにして二つの観察対象を追跡した結果を、図 6.20 に示す。いずれも観察対象がフレームアウトした時刻、そして他の対象に遮蔽された時刻においては認識に失敗している。

この認識結果を V-FES の入力として用い、視覚重畳関係、運動特徴を推定した過程および結果を図 6.21 に示す。グラフに描画された色とメタ情報の対応は図 5.5 に示す通りである。観察 HRP2-OBS3 ではタオルが箱に入れられて運ばれている際に、二つの対象の距離が近く重畳関係が発生し、運動に随伴性が見られることが推定された。一方で観察 HRP2-OBS4 では、初めに対象同士の距離が近く推定されているものの、視覚重畳関係は独立と計算され、その後の運動の随伴性も見られない結果となっている。

さらに V-FES により、観察 HRP2-OBS3 と HRP2-OBS4 に表れる作用と観察対象に表れる道具としての機能を推定させた結果を、それぞれ図 6.22 と図 6.23 に示す。観察 HRP2-OBS3 ではタオル (OBJECT0) が箱 (OBJECT1) に入れられてから棚へ運ばれていること、観察 HRP2-OBS4 ではタオルを机の上に残して、箱のみを棚へ運んでいることが正しく推定されている。

続いて、観察学習させたタスク目的と道具利用法に基づき、箱やタオルを片付けさせる実験をおこなった。実験開始の状況として、初期状態で HRP-2 は机の前におり、机の上には箱とタオルが置いてあるものとした。このとき、観察 HRP2-OBS3 および観察 HRP2-OBS4 の学習結果に基づくタスク計画結果は、各々図 6.24、図 6.25 のようになった。図 6.24 と図 6.25 を実行させた結果が図 6.26 の (A) と (B) である。視覚教示によって、片付けシーンにおける物体投入を含む物体操作が実現可能であることが示された。

ただし、この実験の観察学習においては、タオルを机の上に置くこと自体もタスク目的として扱われている。従って、タスク再現時にタオルが机の上に無かったならば、タオルがある場所まで行ってタオルをつかみ、机の前に来てタオルを置くこ

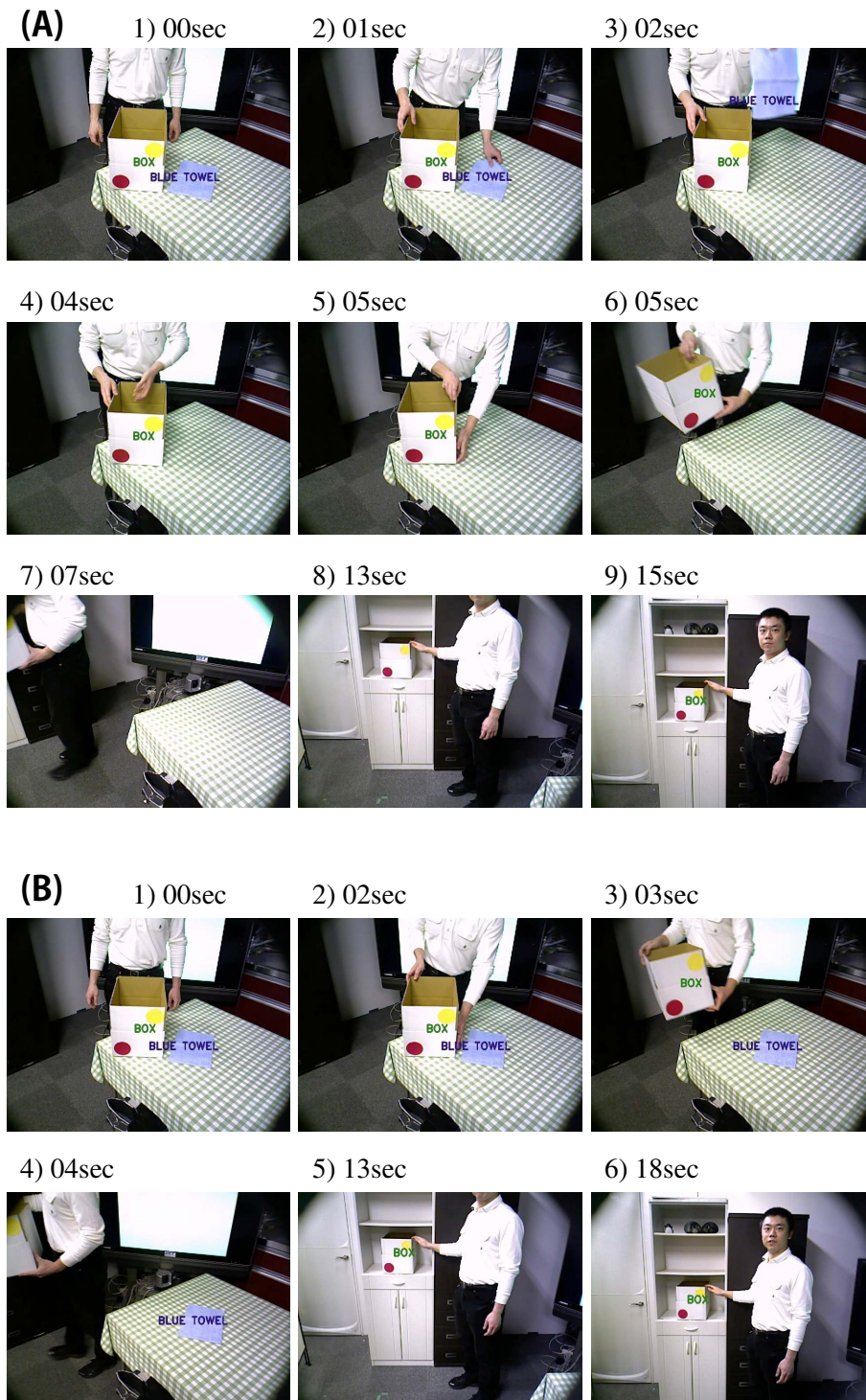


図 6.20: (A) 観察 HRP2-OBS3, (B) 観察 HRP2-OBS4 における観察対象の追跡結果．観察対象の認識に成功したフレームでは，観察対象の名称を重畳表示している（“BLUE TOWEL”は“CLOTH”と同義）

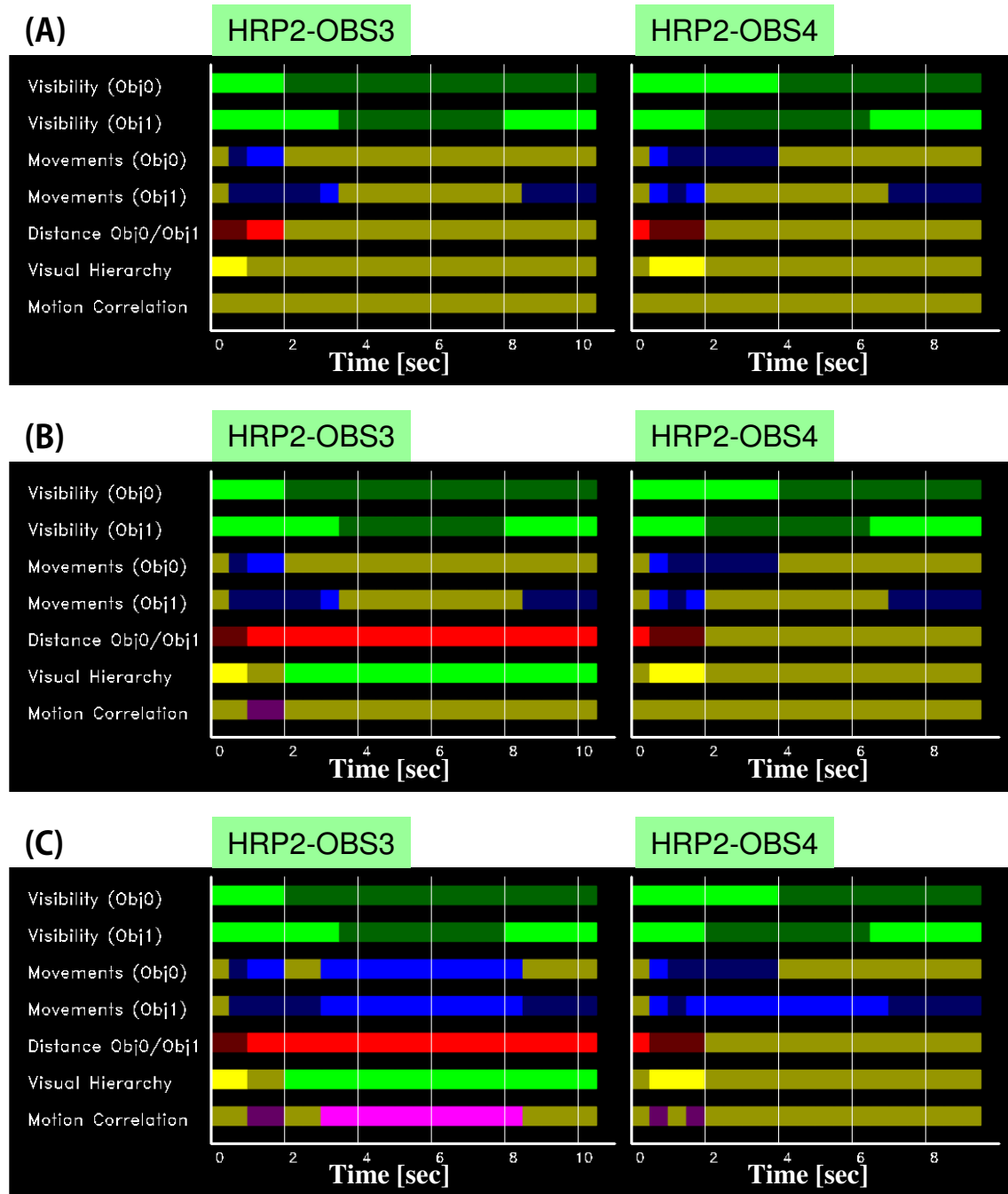


図 6.21: 観察 HRP2-OBS3 および HRP2-OBS4 において、時系列メタ情報を推定した過程と結果。(a) が知識を利用する前、(b) が知識 KNO-1 から知識 KNO-3 までを繰り返し利用した結果、そして (c) が最終的に得られた時系列メタ情報である

```

* OBJECT0 (CLOTH)
AT TIME=0.5[sec], MOVED_FROM PLACE= TABLE,
AT TIME=2[sec], PUT_INTO OBJECT1,
(AT TIME=2.5[sec], START being TRANSPORTED , by OBJECT1)
(AT TIME=8.5[sec], STOP being TRANSPORTED , by OBJECT1)
* OBJECT1 (BOX)
AT TIME=2.5[sec], MOVED_FROM PLACE= TABLE,
AT TIME=8.5[sec], MOVED_TO PLACE= SHELF,
* OBJECT1 CAN BE USED AS STORAGE
* OBJECT1 CAN BE USED AS CONTAINER
;; TASK GOAL:
IN OBJECT0 OBJECT1
AT OBJECT1 SHELF

```

図 6.22: タオルを箱に入れてから棚へ運ぶ観察 HRP2-OBS3 の作用推定・道具機能推定結果

```

* OBJECT0 (CLOTH)
AT TIME=0[sec], MOVED_FROM PLACE= SHELF,
AT TIME=1[sec], MOVED_TO PLACE= TABLE,
* OBJECT1 (BOX)
AT TIME=1[sec], MOVED_FROM PLACE= TABLE,
AT TIME=7.5[sec], MOVED_TO PLACE= SHELF,
;; TASK GOAL:
AT OBJECT0 TABLE
AT OBJECT1 SHELF

```

図 6.23: タオルを机に残したまま棚へ運ぶ観察 HRP2-OBS3 の作用推定・道具機能推定結果

0: (HOLD HRP-2 CLOTH R-ARM TABLE)
1: (PUT-INTO-OBJ HRP-2 CLOTH R-ARM BOX TABLE)
2: (HOLD HRP-2 BOX R-ARM TABLE)
3: (MOVE HRP-2 TABLE SHELF)
4: (PLACE HRP-2 BOX R-ARM SHELF)

図 6.24: 観察 HRP2-OBS3 の学習結果に基づくタスクプランニング結果

0: (HOLD HRP-2 BOX R-ARM TABLE)
1: (MOVE HRP-2 TABLE SHELF)
2: (PLACE HRP-2 BOX R-ARM SHELF)

図 6.25: 観察 HRP2-OBS4 の学習結果に基づくタスクプランニング結果

いう操作を含むタスクが計画されるため、タオルを操作しないということにはならない。この点については、タオルを観察対象と指定しているならば、本観察学習のようにタオルの移動先を記憶することは正しいと言える。タオルを操作してはいけない場合であれば、観察対象として指定せずにおく、つまり人間が動かした対象のみを観察対象とする手法を導入すれば良い。このような手法としては、著者らが提案した、人間の手先と周辺オブティカルフローの相関演算に基づく操作対象領域の切り出し法 [103] が挙げられる。

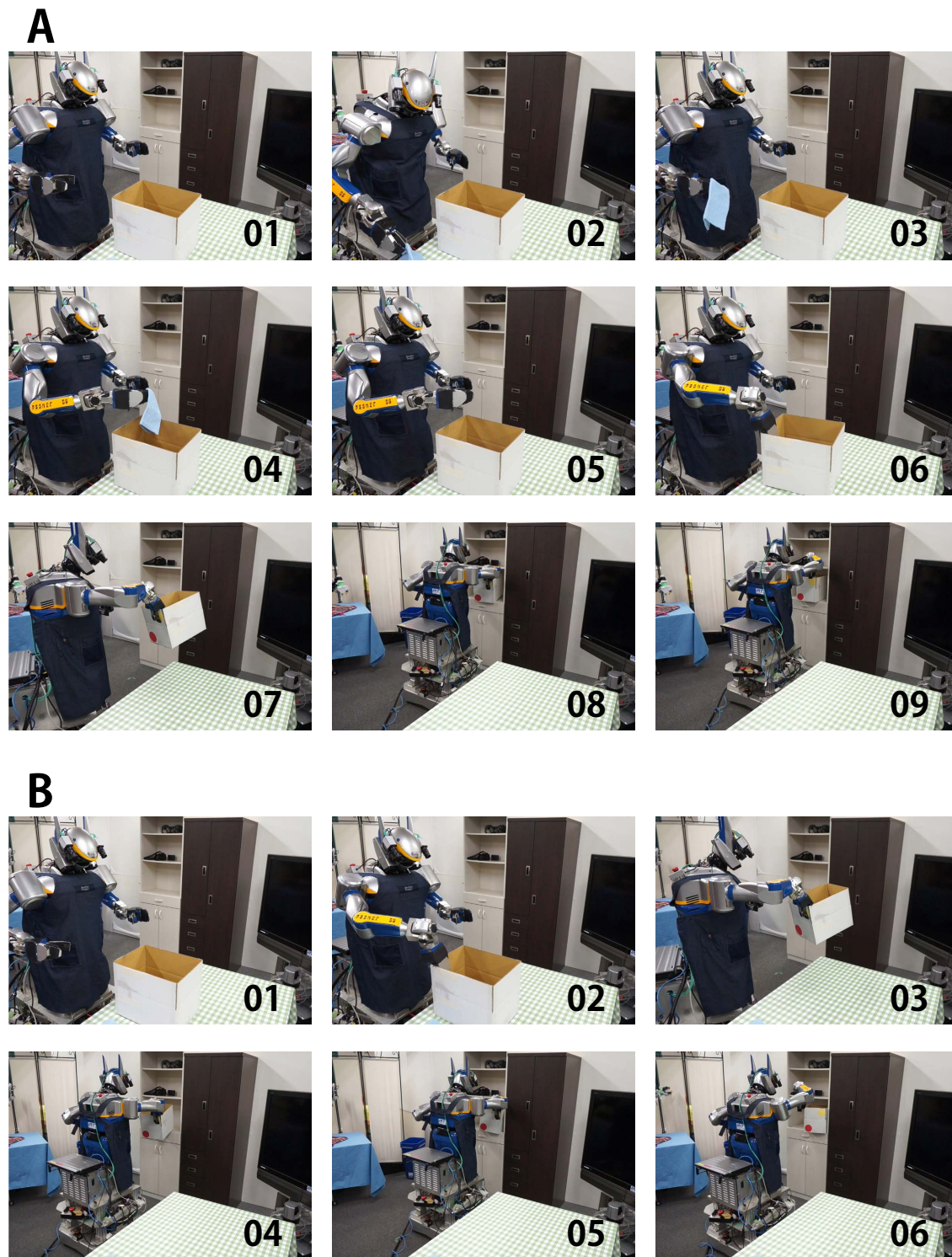


図 6.26: (A) 観察 HRP2-OBS3, (B) 観察 HRP2-OBS4 の学習結果に基づく動作再現結果

6.4.2 洗濯物の分類に関する観察学習とその再現

次に、双腕ロボット HIRO を用いて、柔軟物である洗濯ネットを用いたタスクの観察学習に関する評価をおこなう。本節で観察学習させるタスクは、洗濯物の分類である。洗濯をおこなうにあたっては、洗濯物をその種類によって分類し、何度かに分けて洗濯機を利用する家庭が多い。その際、手袋や一部の下着などは、型くずれや洗濯時の絡みつきを防止するために、洗濯の際に洗濯ネットを用いることが望ましい。このような洗濯物の分類法を簡単に教示できるシステムは、各家庭に応じた洗濯タスク支援に役立つと考えられる。

本実験では靴下・手袋・洗濯ネットを観察対象とし、手袋のみを洗濯ネットに入れてから洗濯カゴに移し、靴下はそのままカゴに移す様子を観察学習させるものとする。V-FES は二つの観察対象同士の関係を推定するものであるが、その学習結果はシンボリックに表現されるため、複数の観察学習結果を以下のように統合することが可能である。

• タスク目的の統合

複数の観察学習結果に矛盾が無ければ、タスク目的のシンボル記述の論理積をとることで、統合されたタスク目的を設定することが可能である。例えば、あるタスク目的 (AND (AT NET BASKET) (IN GLOVE NET)) ともう一つのタスク目的 (AND (AT NET BASKET) (AT SOCK BASKET)) から、(AND (AT NET BASKET) (IN GLOVE NET) (AT SOCK BASKET)) というタスク目的を生成できる。

• 道具利用法の統合

本研究で注目している道具的利用法は、各物体に許される道具としての使い方である。このような道具的利用法は論理和を用いることで、見たことのある利用法を全て利用していくことが可能となる。例えば、ある道具的利用法 (EQ TRUE (CAN-BE-USED-AS-STORAGE? NET)) ともう一つの道具的利用法 (EQ TRUE (CAN-BE-USED-AS-CONTAINER? NET)) は、(AND (EQ TRUE (CAN-BE-USED-AS-STORAGE? NET)) (EQ TRUE (CAN-BE-USED-AS-CONTAINER? NET))) として統合できる。

• 対象の有無と個数によるタスク目的の変更

さらに、タスク再現時に観察学習した対象が存在しない、あるいは観察学習した対

象が複数存在する場合へ拡張する．まず前者へ対応するため，タスク再現時（初期状態）に観察学習した対象を探索し，対象が存在した場合には，その対象に関するタスク目的は学習結果から据え置き，存在しない場合には学習されたタスク目的から省くものとする．例えば，(AND (AT NET BASKET) (IN GLOVE NET) (AT SOCK BASKET)) というタスク目的の学習結果に対し，初期状態が (AND (AT SOCK TABLE) (AT NET TABLE)) であった場合，つまり GLOVE が存在しなかった場合には，(IN GLOVE NET) を削除し，(AND (AT SOCK BASKET) (AT NET BASKET)) を新たなタスク目的として採用する．また，観察対象が二つ存在した場合には，シンボルにインデックスを追加して新たなシンボルを生成し，タスク目的と道具的利用法については観察時と同様のものを付与するのとした．例えば，(AT GLOVE BASKET) というタスク目的の観察学習結果に対し，初期状態が (AND (AT GLOVE TABLE) (AT GLOVE2 TABLE)) として認識された場合，タスク目的を (AND (AT GLOVE BASKET) (AT GLOVE2 BASKET)) として書き換える．

以降，V-FES を用いた観察学習と，上記の手法を用いた統合学習結果から得られる洗濯物分類の動作再現をおこなわせ，有用性を評価するものとした．

観察実験は二回（観察 HIRO-OBS1，観察 HIRO-OBS2）おこなう．観察 HIRO-OBS1 では，洗濯ネット（観察対象 OBJECT0）と靴下（観察対象 OBJECT1）を洗濯カゴへ入れるシーンを観察させる．次に観察 HIRO-OBS2 では，洗濯ネット（観察対象 OBJECT0）へ手袋（観察対象 OBJECT1）を入れてから洗濯カゴへ投入するシーンを観察させる．ただし，観察には HIRO の頭部に搭載した PrimeSense 社の RGBD カメラ “Carmin 1.09” を用いるものとした．

衣類の視覚分類に関しては，折れ重なりやしわの状態に着目して種類を識別する手法 [114]，や，テクスチャを含む特徴点を学習し分類に用いようとする取り組み [87] などがある．また，衣類の形状認識についても，予め学習したオプティカルフローとのマッチングをとる手法 [77] 等がある．本研究の提案手法も，このような視覚分類手法との併用が可能だと考えられるが，本実験では物体の移動と包含関係の変化の学習性能について評価することが目的であるため，簡易的に対象の色を用いて対象を識別させるものとした．基本的に前章で述べ，前実験でタオルの認識に用いた手法と同一であるが，色による対象領域の特定の前に，距離画像のフレーム差分から，机や洗濯カゴ等の固定環境上の領域を除外するものとした．洗濯ネットに物を

投入する場面について、視覚重畳関係と運動の随伴性から物体の上下・包含関係を含む状態変化を推定する V-FES を用いることで、洗濯ネットの複雑な形状を推定することなく、物の投入を認識させることができると期待できる。

図 6.27 および図 6.28 に示した観察 HIRO-OBS1 と観察 HIRO-OBS2 の観察対象の追跡結果を入力として、V-FES を用いて視覚重畳関係、運動特徴を推定した過程と結果を図 6.29 に示す。ただし、図中 (A) が知識を利用する前の時系列メタ情報、(B) が知識 KNO-1 から知識 KNO-3 までを繰り返し利用した結果、(C) が全ての知識を利用した結果であり、グラフに描画された色とメタ情報の対応は図 5.5 に示す通りである。観察 HIRO-OBS1 では、最後に対象同士の距離が変化しているものの、視覚重畳関係は発生せず、運動に随伴性も存在しないことが推定されている。一方で観察 HIRO-OBS2 では、手袋を洗濯ネットに入れた後に洗濯ネットが手袋を隠すような片側遮蔽関係が発生し、さらにネットを洗濯カゴへ移す際に運動の随伴性が存在することが正しく推定されている。さらに V-FES により、観察中に表れる作用と、観察対象に表れる道具としての機能を推定した結果が図 6.30 および図 6.31 である。観察 HIRO-OBS1 では、洗濯ネット (OBJECT0) と靴下 (OBJECT1) が独立に洗濯ネット (BASKET) へ移動されていることが推定されている。一方で観察 HIRO-OBS2 では、手袋 (OBJECT1) を洗濯ネット (OBJECT0) へ入れてから (3[sec])、両方をまとめて洗濯カゴ (BASKET) へ投入していることが正しく推定されている。

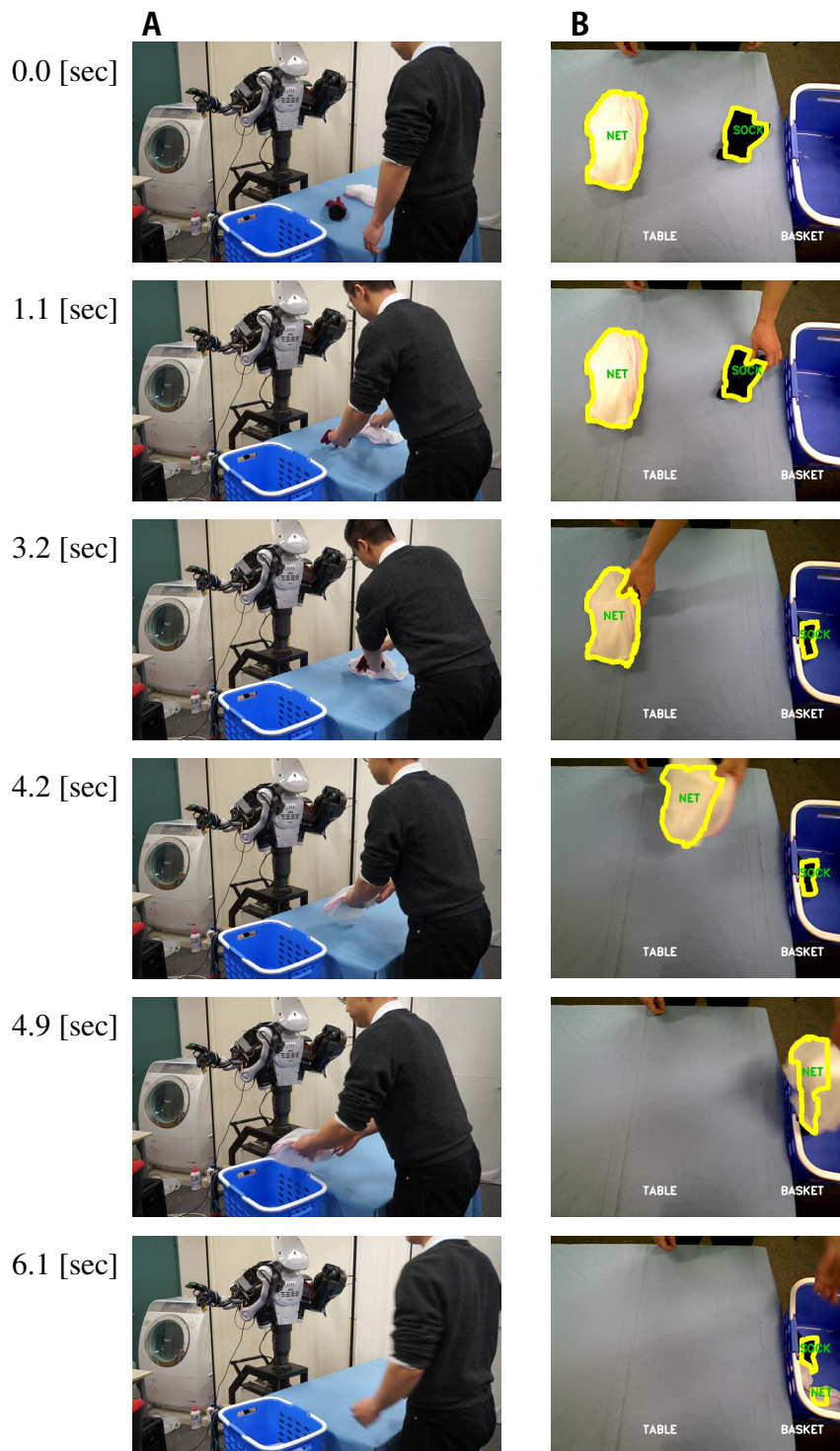


図 6.27: 観察 HIRO-OBS1 (洗濯ネットに入れずに洗濯カゴへ移す場合) の観察結果．外部カメラから撮影したものが A，HIRO のカメラを用いて洗濯物を追跡した結果が B である．認識された対象は黄線で囲まれ，緑文字で名称を表示されている

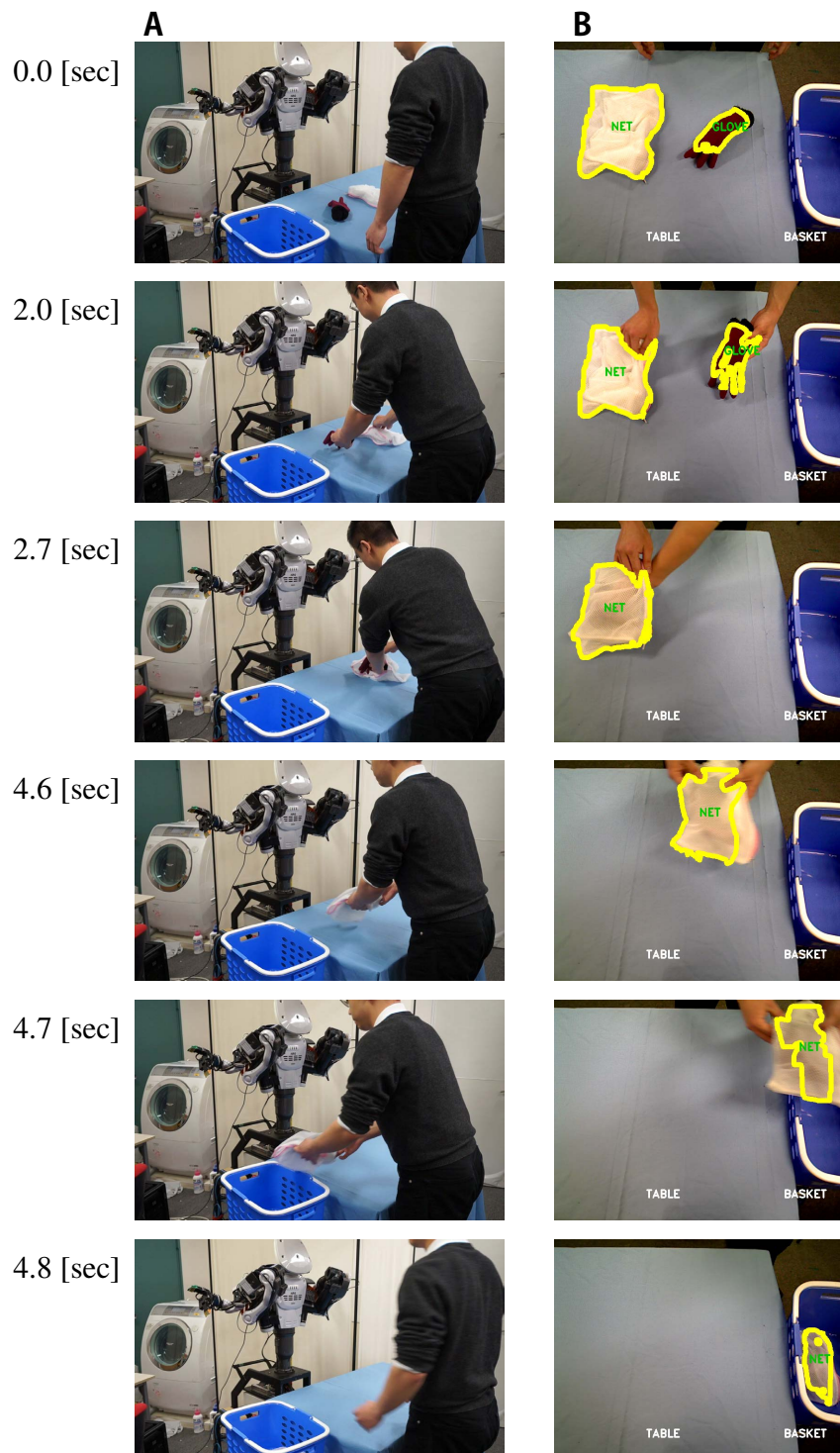


図 6.28: 観察 HIRO-OBS2 (洗濯ネットに入れてから洗濯カゴへ移す場合) の観察結果．外部カメラから撮影したものが A，HIRO のカメラを用いて洗濯物を追跡した結果が B である．認識された対象は黄線で囲まれ，緑文字で名称を表示されている

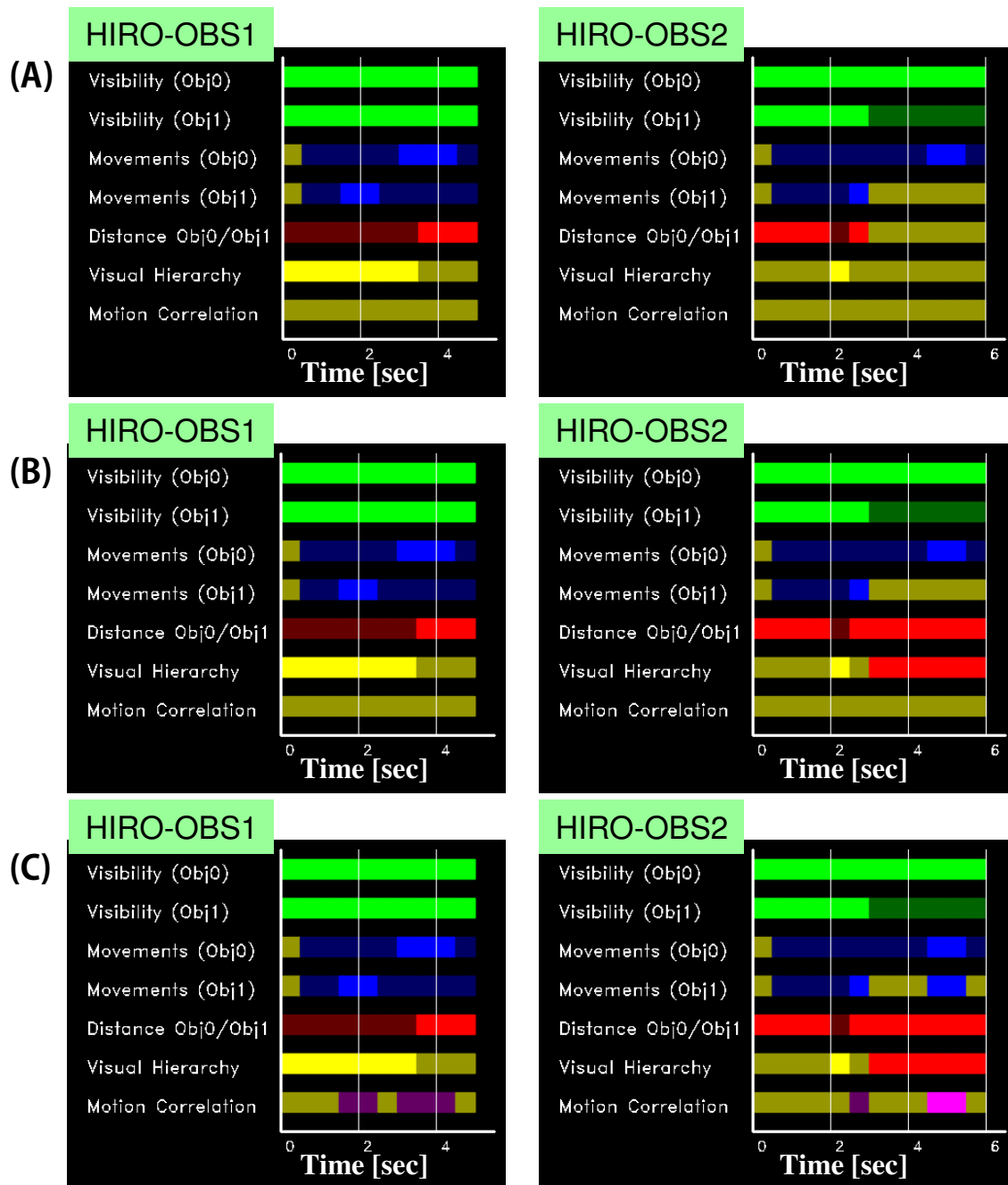


図 6.29: 観察 HIRO-OBS1 および HIRO-OBS2 において，時系列メタ情報を推定した過程と結果．(a) が知識を利用する前，(b) が知識 KNO-1 から知識 KNO-3 までを繰り返し利用した結果，そして (c) が最終的に得られた時系列メタ情報である


```
* OBJECT0 (NET)
AT TIME=2.5[sec], MOVED_FROM PLACE= TABLE,
AT TIME=4.5[sec], MOVED_TO PLACE= BASKET,
* OBJECT1 (SOCK)
AT TIME=1[sec], MOVED_FROM PLACE= TABLE,
AT TIME=2.5[sec], MOVED_TO PLACE= BASKET,
;; TASK GOAL:
AT OBJECT0 BASKET
AT OBJECT1 BASKET
```

図 6.30: 靴下と洗濯ネットを洗濯カゴに入れる観察 HIRO-OBS1 の作用推定・道具機能推定結果．OBJECT0 が洗濯ネット，OBJECT1 が靴下である

```
* OBJECT0 (NET)
AT TIME=4[sec], MOVED_FROM PLACE= TABLE,
AT TIME=5.5[sec], MOVED_TO PLACE= BASKET,
* OBJECT1 (GLOVE)
AT TIME=2[sec], MOVED_FROM PLACE= TABLE,
AT TIME=3[sec], PUT_INTO OBJECT0,
(AT TIME=4[sec], START being TRANSPORTED , by OBJECT0)
(AT TIME=5.5[sec], STOP being TRANSPORTED , by OBJECT0)
* OBJECT0 CAN BE USED AS STORAGE
* OBJECT0 CAN BE USED AS CONTAINER
;; TASK GOAL:
AT OBJECT0 BASKET
IN OBJECT1 OBJECT0
```

図 6.31: 手袋を洗濯ネットに入れてカゴに入れる観察 HIRO-OBS2 の作用推定・道具機能推定結果．OBJECT0 が洗濯ネット，OBJECT1 が手袋である

```

1 (:action PUT-INTO-OBJ
2     :parameters '((?robot_ robot) (?object_ object)
3                 (?arm_ arm)
4                 (?object_into_ object)
5                 (?arm_2 arm)
6                 )
7     :precondition '((hold? ?robot_ ?object_ ?arm_)
8                   (hold? ?robot_ ?object_into_ ?arm_2)
9                   (can-be-used-as-storage? ?object_into_)
10                  (near? ?robot_ table) ;; laundry
11                  (clear? ?object_into_)
12                  )
13     :effect '((handempty? ?robot_ ?arm_)
14              (not (hold? ?robot_ ?object_ ?arm_))
15              (in? ?object_ ?object_into_)
16              (increase (total-cost) 1)
17              (not (clear? ?object_into_))
18              )
19    ))

```

図 6.32: 柔軟性を有する袋に物体を投入するための PDDL action 定義

次に、観察学習させたタスク目的と道具利用法に基づいて、洗濯物を分類させる実験をおこなった。実験開始の状況として、初期状態で HIRO は机へ向かっており、机の上には靴下や手袋があり、左手で洗濯ネットを掴んでいるものとする。タスク再現実験 HIRO-EXP1 では、机の上には手袋と靴下がある。実験 HIRO-EXP2 では靴下二枚と手袋があり、実験 HIRO-EXP3 では靴下のみがあるものとした。

HIRO による洗濯ネットの開口動作と物体投入動作は、距離画像のエッジから開口部を認識し、操作中の自らの手の位置を用いて物体投入のための操作点を決定する手法 [97] を用いる。この手法によると、柔軟性を有する袋の中に物体を投入する際には、片手で袋を持ち、形状を保ちながらもう一方の手で物体を投入する戦略が有用である。そこで本実験でもこの戦略を利用するために、洗濯ネットへの物体投入時は洗濯ネットを支えているよう、タスクプランニング時の物体投入を表す action を図 6.32 のように記述した。また、物体のつかみ上げ (HOLD) 時には、初期状態認識の際の物体の位置認識結果を用いるものとした。

実験 HIRO-EXP1 から実験 HIRO-EXP3 の，初期状態における各操作対象の認識結果を図 6.36 に示す．本節はじめで述べた手法により，新たなタスク目的を生成した結果，それぞれ以下ようになった．

実験 HIRO-EXP1 のタスク目的

(AND (AT NET BASKET) (IN GLOVE NET) (AT SOCK BASKET))

実験 HIRO-EXP2 のタスク目的

(AND (AT NET BASKET) (IN GLOVE NET) (AT SOCK BASKET) (AT SOCK2 BASKET))

実験 HIRO-EXP3 のタスク目的

(AND (AT NET BASKET) (AT SOCK BASKET))

結果，タスクプランニングの結果は HIRO-EXP1 から HIRO-EXP3 までそれぞれ図 6.33 から図 6.35 のようになり，図 6.37 から図 6.39 に示すように，期待したような洗濯物の分類と洗濯カゴへの投入に成功した．

本実験においては，物体の存在可否に関わる一連の動作の変更はタスクレベルでの動作変更によって，また視野内の物体の移動に関わる各動作の小変更は認識・動作レベルでの戦略変更によって吸収することにより，観察学習時とは異なる環境においても分類操作が可能であった．柔軟で形状の変化する洗濯ネットに対しても，包含関係の変化を含むタスク学習とその再現が可能であり，生活支援タスク例における本手法の有用性が示されたと言える．ただし，本節で述べたタスク目的の統合と修正手法に関しては注意が必要である．対象が存在する場合に，対象同士の間関係を観察時に近づけることが重要である状況，つまり片付けや本実験のような分類タスクの場合には，本手法はそのまま有効である．しかし，複数の対象同士の特定の関係が必ず必要であるような状況では，このようなタスク遂行は適切ではなくなる．このような状況では，タスク実行を停止し，足りない操作対象について人間へ確認するのが良いと考えられるため，将来的にはタスク目的の修正をすべきではない場合に，それを教えることの可能なシステムとしていくべきだと考えられる．

```
0: (HOLD HIRO GLOVE R-ARM TABLE)
1: (PUT-INTO-OBJ HIRO GLOVE R-ARM NET L-ARM)
2: (HOLD HIRO SOCK R-ARM TABLE)
3: (MOVE HIRO TABLE BASKET)
4: (PLACE HIRO SOCK R-ARM BASKET)
5: (PLACE HIRO NET L-ARM BASKET)
```

図 6.33: 観察 HIRO-OBS1, 2 の学習重ね合わせ結果に基づく, 実験 HIRO-EXP1 のタスクプランニング結果

```
0: (HOLD HIRO GLOVE R-ARM TABLE)
1: (PUT-INTO-OBJ HIRO GLOVE R-ARM NET L-ARM)
2: (HOLD HIRO SOCK2 R-ARM TABLE)
3: (MOVE HIRO TABLE BASKET)
4: (PLACE HIRO SOCK2 R-ARM BASKET)
5: (PLACE HIRO NET L-ARM BASKET)
6: (MOVE HIRO BASKET TABLE)
7: (HOLD HIRO SOCK R-ARM TABLE)
8: (MOVE HIRO TABLE BASKET)
9: (PLACE HIRO SOCK R-ARM BASKET)
```

図 6.34: 観察 HIRO-OBS1, 2 の学習重ね合わせ結果に基づく, 実験 HIRO-EXP2 のタスクプランニング結果

```
0: (HOLD HIRO SOCK R-ARM TABLE)
1: (MOVE HIRO TABLE BASKET)
2: (PLACE HIRO SOCK R-ARM BASKET)
3: (PLACE HIRO NET L-ARM BASKET)
```

図 6.35: 観察 HIRO-OBS1, 2 の学習重ね合わせ結果に基づく, 実験 HIRO-EXP3 のタスクプランニング結果

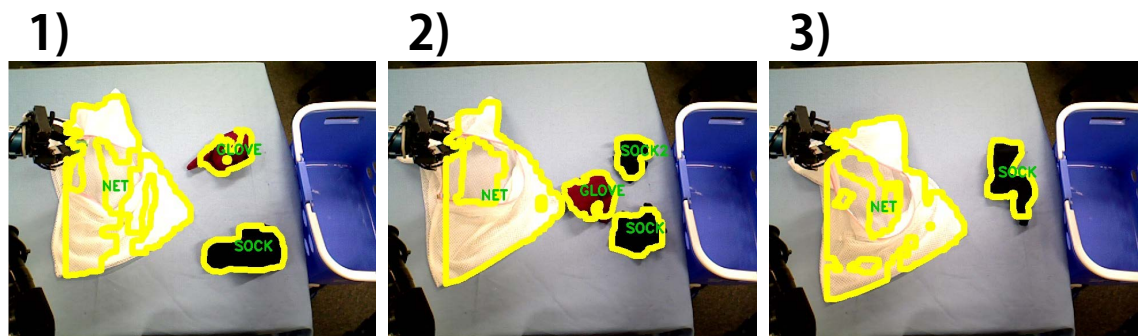


図 6.36: 観察 HIRO-OBS1, 2 の学習重ね合わせ結果に基づくタスク再現実験 HIRO-EXP1 から HIRO-EXP3 の初期状態認識結果．認識された対象は黄線で囲まれ，緑文字で名称を表示されている．1) が HIRO-EXP1 ，2) が HIRO-EXP2 ，3) が HIRO-EXP3 の初期状態における操作対象認識結果である

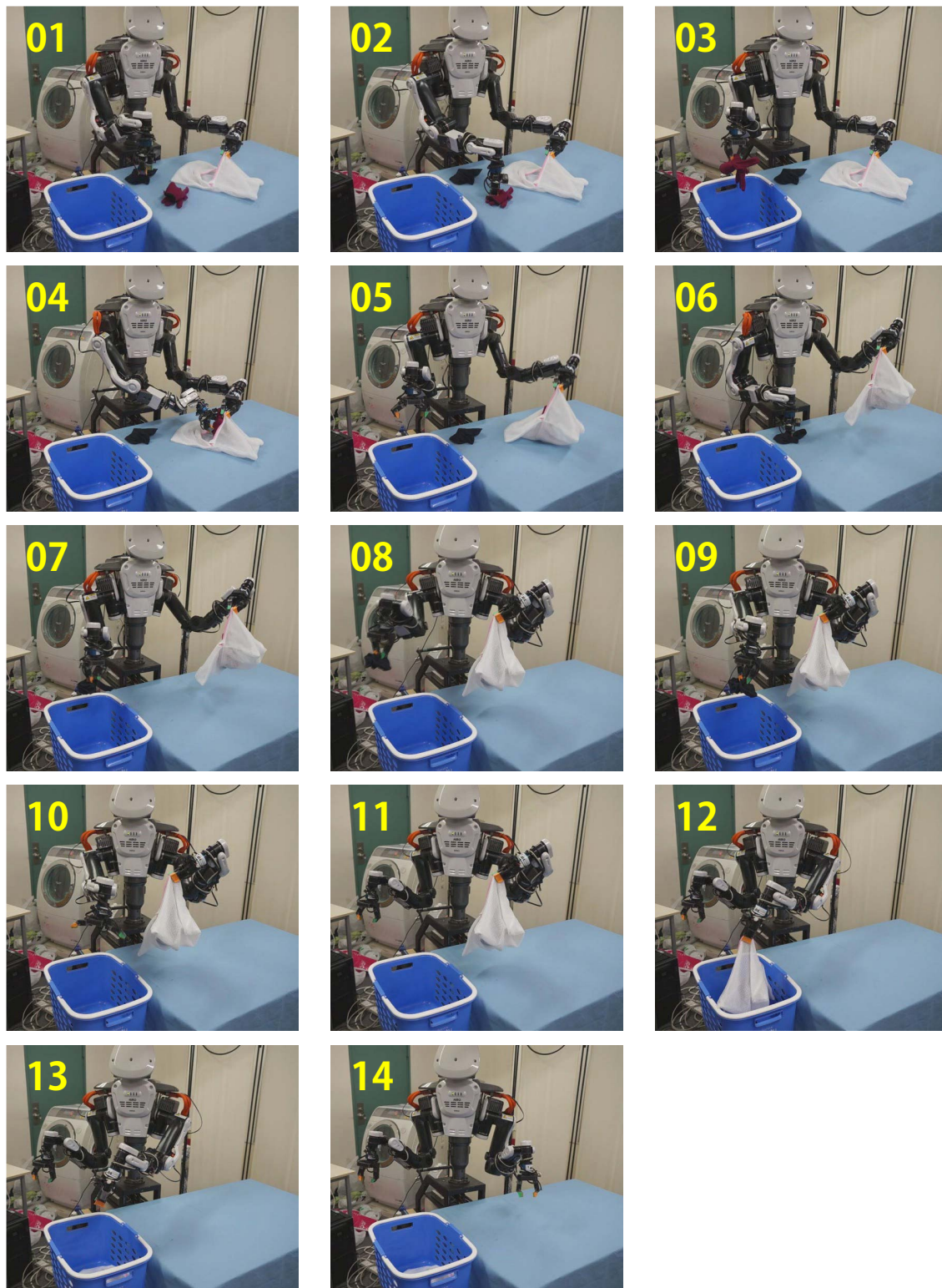


図 6.37: 観察 HIRO-OBS1, 2 の学習重ね合わせ結果に基づくタスク再現実験 HIRO-EXP1 の結果

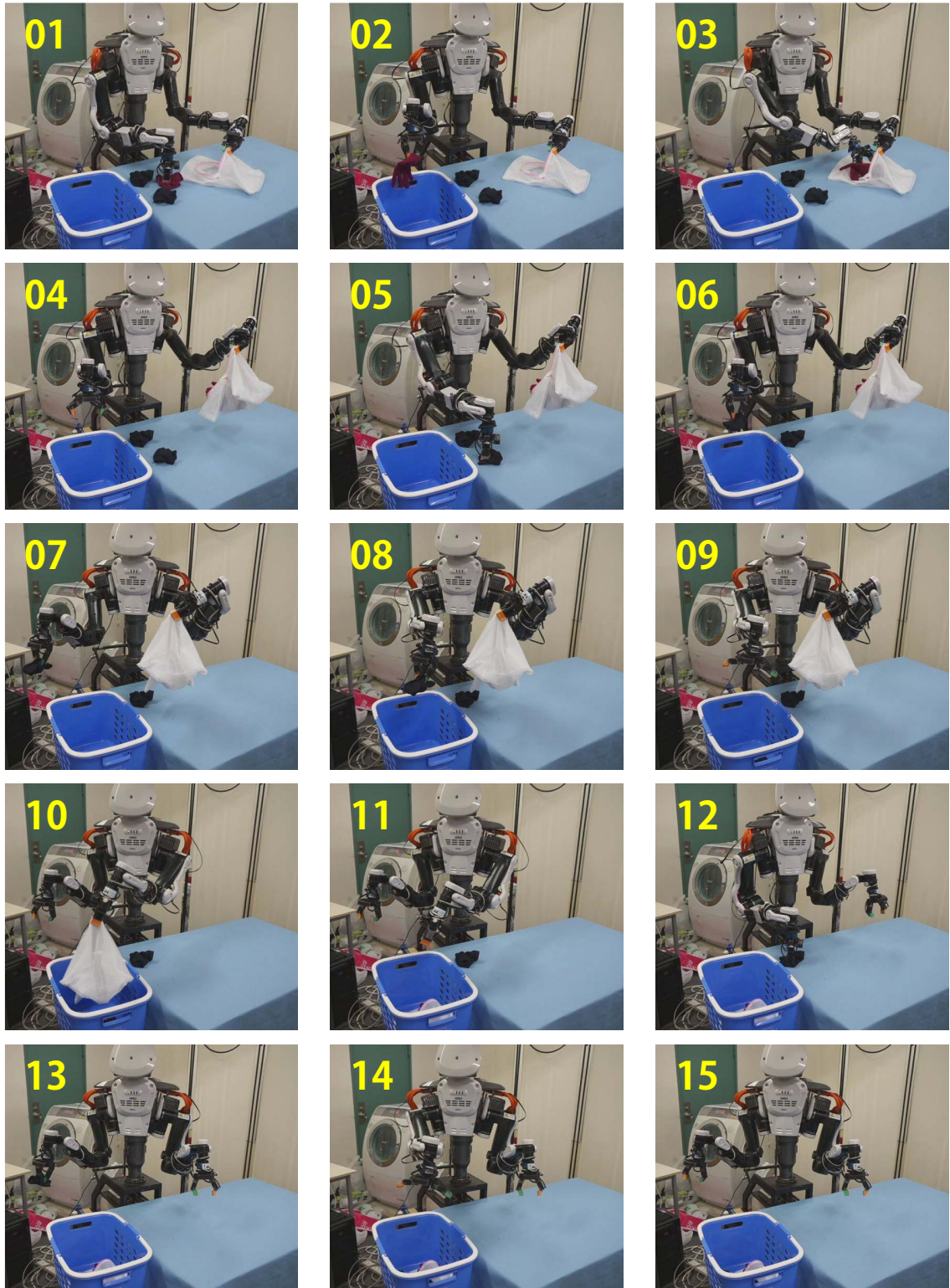


図 6.38: 観察 HIRO-OBS1, 2 の学習重ねあわせ結果に基づくタスク再現実験 HIRO-EXP2 の結果

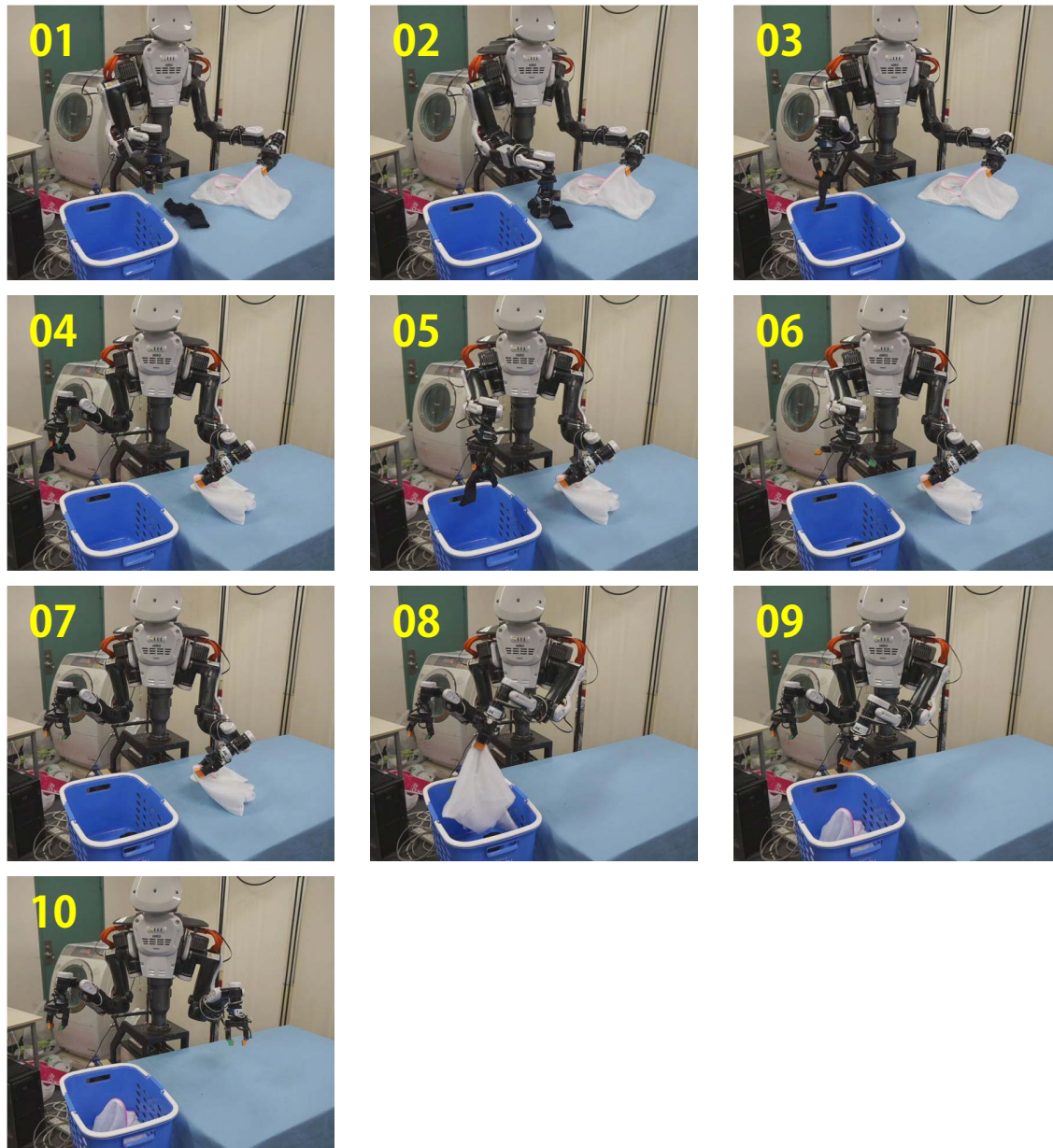


図 6.39: 観察 HIRO-OBS1, 2 の学習重ねあわせ結果に基づくタスク再現実験 HIRO-EXP3 の結果

6.5 おわりに

本章では、前章までの道具を用いたタスクの観察学習結果に基づき、実行時のロボットの体と環境に合わせ、各物体に許される道具的利用法のみを用いて、学習したタスク目的を再現する方法について述べた。本章までの観察学習手法およびタスク再現法を接続することで、人間の教示した、物体の上下・包含関係の変化を伴う「載せる」「入れる」といった動作を含むタスクを書き下し、ロボット自身がそのタスクを再現していくことが可能であることが示された。さらに、この観察学習法およびタスク再現手法が、実環境での片付けや洗濯物の分類といった生活支援シーンにおいても有用であることを、実証的に明らかにした。これにより、2章で指摘した「視覚重畳関係を用いた物体同士の関係推定を用いたタスク学習の可能性」が確認され、「生活支援タスクの学習・実行に際しての有用性」が実証された。

今回の各実験では、ロボットそのものの移動にかかるコストをマニピュレーションのコストよりも大きく設定したため、積極的に道具を使ってから移動するような戦略が高い確率で選択された。計測はおこなっていないものの、ロボットそのものの質量は十分に重いため、タスク全体にかかわる消費電力は道具を利用しない場合に比べ、抑えられているだろうと考えられる。さらに、複数の対象を運ぶべき場所が遠くなればなるほど、この傾向は大きくなる。消費電力を正しく予測することができれば、その消費電力を、予測結果をコストに反映して消費電力を最小にするようなプランニングをおこなうことができる。ただし実際には、コストとして何を用いるかは、臨機応変に決定する枠組みが必要となると考えられる。例えば、タスクにかかる時間、人間が横で見ている場合の騒音や見た目の煩雑などを計算することができれば、家庭内でより親和性の高いタスク再現が可能であるだろうと考える。

また本章では、特に前章で推定したタスク目的と道具的利用法の利用可能性について評価した。4章でその推定法について述べた、道具が物体へ突入するような作用を伴うタスク目的の実現に関しては、突入する部位あるいは切断面と切断面の関係も重要となるであろうから、突入する作用が推定された場合には、そのような動作に必要な対象目標レベルのパラメタをタスク目的に加えて記憶し、突入や切断をおこなう action の実行時に利用することで再現が可能となると考えられる。

このような対象目標レベルのパラメタの数をインタラクティブに増やし設定する

ことが可能となれば、前章から本章で述べたような上下・包含関係を変化させるタスク学習についても、引き出しの中のどの部分にしまうべきか等のより詳細な学習と再現が可能となっていくと考えている。

道具の選択に関しては、本手法のような決定法によれば、一度見たことのある道具利用法のみが活用されることとなる。これは、家庭内では他の対象に作用すべき道具は各家庭によって独自に限られており、その規則を学習すべきだという見解に基づくものであった。また、前章および本章での手法によると、道具の機能は操作対象によらず利用可能となっている。ただし実際の家庭では、調理に用いるものは調理のみに用い、掃除に用いるものは掃除のみに用いるなど、同様の機能を有する道具についても、文化・衛生上の観念から使い分けがなされている状況が多い。この点を提案手法へ組み込むことを考えると、道具の機能をより詳細に分割すれば、操作対象による使い分けが実現可能となると考えられる。道具と操作対象について全て記憶することも考えられるが、実際には、道具は食品用の道具、掃除用の道具、文房具などと分類され、それぞれの利用場面ごとに同様に操作できる対象が多く存在すると考えられる。道具の機能の分類のうち操作対象を限る性質に関して、多くの操作対象と道具の組み合わせについてデータを収集し、その家庭の文化的な道具利用法を学習して新規操作対象への道具選択が可能としていくことが、展望して挙げられる。

第7章

結論

7.1 各章のまとめと結論

本論文は、生活支援ロボットが道具を用いておこなうべき各家庭に固有の家事について、人間を観察し、その目的と観察対象の道具としての使い方を推定・記憶することで、家事をロボットが可能な形で再現していく際の問題点と、複数の観察対象あるいは複数の操作対象同士の遮蔽関係に着目した解決法およびその有用性について明らかにするものであった。

第1章序論では、各家庭で生活支援ロボットへ人の手助けをおこなわせる際に必要となる、各家庭固有の家事方法を指示するための手法について比較検討し、視覚教示の有用性を指摘するとともに、本論文の構成を述べた。

第2章作用・機能推定に基づく道具利用タスク学習では、まず従来のロボットによる物体操作の視覚学習研究と生活支援ロボットによる道具操作についてまとめ、観察対象や操作対象が薄い、あるいは柔軟性を有する場合に、対象同士の上下・包含・突入関係の変化を、対象に遮蔽が発生している場合においても推定することの重要性和難しさを指摘した。次に観察対象や操作対象同士の遮蔽関係である視覚重畳関係を陽に用いることが、対象の観測や対象同士の関係の推定に有用である可能性を指摘した。重畳関係は、一般的なカメラで撮影した画像中で、複数の対象の表面を表す領域同士が、互いに隠しあるいは隠される関係である。人間と類似した観察者視点を与えた場合、二次元相互重畳関係を用いることで三次元的な包含・突入関係を、二次元片側遮蔽関係を用いることで上下関係を、そして運動情報を併用することで、一方が完全に遮蔽されている上下・包含関係を推定できる可能性について述べた。

第3章道具を用いた生活支援タスクを学習するロボットシステムでは、日用道具を用いて生活支援をおこなうロボットへの要求機能を述べた。そして、人間と同様の日用道具を用いるマニピュレーション能力と人間並みの視力を有するロボットの一実現例として、実験に用いる等身大ヒューマノイドのハードウェア及びソフトウェア構成法について述べた。さらに広視野視覚と高視力視覚を併用し、連続視差フィルタを用いて人間の上半身を追跡しながら、手元へ高視力視覚を誘導する実験では、ロボットが自律的に道具と操作対象を高い解像度で観測することが可能であることを示した。また、料理のよそいつけ等の、道具を用いた間接的な物体操作を含む基

礎的な調理実験により、道具利用タスクに必要な視覚処理と失敗復帰の重要性について考察した。

第4章空間遮蔽近傍情報を用いた見えない道具の状態推定では、近接あるいは接触しあう道具と操作対象を観察する際、あるいはロボット自身が道具を利用する状況を想定し、強い遮蔽関係が生じる状況下でも道具と操作対象を追跡し続け、重畳関係を推定することが可能か、そして重畳関係を活用した道具と操作対象の関係推定および、それに基づくロボットの動作学習が可能か否かを検証し、遮蔽された領域の空間的な近傍の情報を用いることで、いずれも可能だという知見を得た。この章では道具と操作対象を追跡対象とし、粒子フィルタを用いた各領域の追跡をおこなっている。テクスチャの少ない日用道具等に対応し、かつ非遮蔽予測領域のみを用いた尤度計算法を導入することで、追跡対象同士に大幅な遮蔽関係が生じた場合へも対応した。非遮蔽予測領域は、追跡対象領域のうち、次フレームにも見えているであろう領域のことであり、現フレームにおいて他方の対象との間に重なり関係が生じていない領域として計算する。さらに、二つの追跡対象の重なり合う領域について、いずれの対象を手前とした方が尤もらしいかを輝度の差を利用して計算し、重畳関係を決定する。操作中の包丁と果物を同時追跡しながら操作分類をおこなう実験では、決定した重畳関係と運動の随伴性情報を併用することで、操作対象を上に乗せて動かす、操作対象に切り込むという上下・突入関係の変化を起こす操作の推定がおこなえることを実証した。さらにロボット自身による料理のよそいつけとファイリング作業という、操作対象の柔軟性や摩擦力の変動により、成功のための動作パラメタが毎回異なり、失敗復帰が必須となる道具利用実験をおこなった。ここでは提案手法を用いて道具と操作対象の上下関係や包含関係を推定し、その関係を用いて観測しにくい方向の動作パラメタを勾配的に変化させることで、操作を成功に導くことが可能であることを示した。

第5章遮蔽時刻近傍情報を用いた見えない作用推定に基づく学習法では、袋への収納や片付け等の観察において、複数の観察対象がフレームアウトや互いの遮蔽関係により全く見えなくなる状況を想定し、そのような場合でも観察対象同士の上下・包含関係の変化を推定することが可能か否かを検証し、遮蔽された時刻の近傍情報を用いることで可能だという知見を得た。本章では偽陽性の低い視覚探索手法によって、観察対象が認識可能な時刻のみで位置を特定した認識結果を入力とし、観察対

象の動きだけでなく、他方の観察対象に載った、あるいは入った、そしてその状態でどこかへ動かされたという関係の変化を推定する認識器を構成し評価した。関係の推定には観察対象同士の重畳関係と運動特徴を用いるが、観察対象が認識されていない時刻にはこれらを直接推定することはできない。そこで、観察対象の視認性の有無を含む五つの視覚的手がかりを定義し、各時刻において推定すると共に、計算不可能な時刻の手がかりについて、前後の時刻の他の手がかりを用いて推定するための幾つかの知識を用いる手法を提案した。色による領域推定や改良 LINE-MOD 法による対象追跡結果を入力する実験では、箱だけでなく、袋のような形状の変化する物体を対象とした観察学習においても、複数の物体同士の上下・包含関係の変化が推定可能であり、その変化を用いてタスク目的と、観察対象を道具として利用して良いという利用可能情報が推定可能であることを示した。

第6章作用・機能推定結果に基づくロボットの身体に合わせたタスク再現では、前章までの手法により推定した家事タスクの目的と物体の利用可能情報という知識が、ロボットがタスク再現時の状況やロボット自身の身体性能に合わせて、道具を用いた家事タスクをおこなう際に有用であることを示した。本章では、教示された家事タスクを、作業目標レベルで再現可能とするために、物体同士の上下・包含関係と道具の利用法を表す述語を定義し、STRIPS タイプの因果関係記述とタスクプランナを用いる手法を導入した。双腕・単腕ロボットによる食器片付け実験により、ロボットの身体に合わせたタスク計画と実行が可能であること、そして洗濯ネットに入れるべきものを学習し、教わった通りに洗濯物を分類していく実験により、本手法によるシンボリックなタスク学習結果を重ねあわせることの可能性と有用性を示した。

以上、本論文は生活支援ロボットが道具を利用する人の行動の観察からその行動タスクを学習するために、複数の観察対象の互いの遮蔽関係や随伴関係を空間的近傍および近傍時刻の観察結果に基づいて推定し、タスク中の物体の上下・包含・突入関係を変化させる行為を視覚学習させる方法に関する研究をまとめたものであった。その結果、調理道具による切る・載せる・運ぶなどの基本操作、箱や袋への収納操作、視野から見えなくなる操作などを含む家事タスクの観察学習を行い、観察記述と状態遷移プランナを利用し、道具利用行動の成否判定と失敗復帰を含むタスク行動を実現できることを実験とともに示した。

7.2 将来の展望

本稿で検証に用いたタスク学習システムは、それ自身が生活支援タスクの学習および遂行へ有益なものであったが、そのまま扱える範囲には限界が存在する。これは、システムが主に作業目標レベルでのタスク学習をするよう設計されており、そのレベルで表現できる事象は限られるからである。例えば、道具利用学習においては軌道レベルでの道具の動かし方の学習も重要となる。ただし、本稿で提案した時空間遮蔽近傍特徴を用いる手法は、特に空間遮蔽近傍特徴を用いる手法について、遮蔽下にあっても観察対象の軌道が得られるものであった。そして空間遮蔽近傍特徴を用いる手法、および遮蔽時刻近傍特徴を用いる手法はいずれも、遮蔽関係を変化させるような作用の推定と分節化に有用なものであった。従ってこれらは従来より研究されている軌道レベルの視覚教示、あるいは作業目標レベルの視覚教示へも、学習のための入力軌道として、あるいは学習すべき操作の分節化のために、活用可能だと考えられる。特に、教示者が観察対象をきちんと見せることなく、観察対象が完全に遮蔽下にあるような観察学習をおこなうということは、遮蔽下での観察対象の動きはそれほど重要でないことを示している可能性がある。この場合、軌道レベルでの学習においては、観察対象が見えなくなる前後で操作を分割し、見せられている部分を重点的に学習していくことが可能であろう。時空間近傍特徴に基づく観察対象の軌道獲得結果を、本稿で検証に用いたタスク学習システムとは異なるレベルでの学習へ用いることで、より複雑な作業の観察学習が可能となっていくと考えられる。

また、対象目標レベルの記述法についても、本システムでは近接する対象同士の関係は重要な関係として全て記憶・再現するものとなっている。ただし、例えばゴミ箱に二回ゴミを投入した場合のゴミとゴミの関係等は、記憶と再現が不要な関係である。展望としては、本手法に基づき、認識され得る限りの観察対象間の関係を推定した後、インタラクティブに不要な関係を無くすよう指示していくシステム、そしてその指示結果を蓄積して、自律的に不要な関係を除去していくシステムへと発展していくべきだと考えている。ただし、そのようなシステムへ発展していく中でも、物体同士の関係の遮蔽性は本質的に重要な手がかりとなるはずである。観察対象が見えないことを認識できることは、観察学習において不要な観察対象あるいは

観察対象同士の関係を見せていないことの認識の手がかりと成り得ると考えられるからである。ゴミがゴミ箱へ入れられた場合に、改めてゴミ箱の中身を見せられることがなければ、ゴミ箱の中でのゴミとゴミ同士の関係は提案システムではタスク目的として推定されることがない。これは、オープンフレームな観察対象同士の関係推定を一段階狭めることに貢献すると考えられる。

また、本論文での実験において観察すべき対象は、人間が画像上で指示すること、あるいは探索すべき特徴を指定することにより決定している。これをロボットに自律的におこなわせるための手法としては、ジェスチャにより指示する、あるいは人の手の動きに随伴して動く対象として認識する手法 [103] 等がある。調理作業など、重要な観察対象が時々刻々変化していく場合には課題が残るが、人間が操作している対象、あるいは注目してきた対象に遮蔽関係が生じたことを検知し、遮蔽関係を生じさせた対象として決定することも展望として挙げられる。この考え方はゴミ箱や机等の環境に固定された対象に関しても適用可能であると考えられる。さらに人の手をも観察対象とすれば、人の手と、その先の操作対象との間に生じている相互作用は、本論文で示した重畳関係を基軸とした手法を用いることで推定していくことが可能であると考えている。

さらに本研究の手法では、三次元形状を把握することが難しく、特に裏側の生地 の形状を知ることが遮蔽により困難であるような布地についても、布地と他の物体との上下・包含関係を推定することが可能であった。このことは、作用推定に有用なだけでなく、逆にこの機能情報を用いて、およその布地の形状情報を推定することの可能性を示すものだと考えている。

視覚重畳関係を有効に利用できるような視点の自律決定に関しても、将来の展望として挙げられる。包丁や下敷きなどの平面状の道具は、その平面内に視点を置くと、視覚で捉えられる領域は小さく、上下・包含・突入関係の推定に有用であった片側遮蔽関係や相互重畳関係の推定が難しい。これに対し、観察時には道具の平面から離れた視点をとって観察する、ロボット自身の操作実行時には、視線方向を考慮しながら道具を動かす方向を決定する等をおこなえば、道具と操作対象同士の関係推定をおこない易くなると考えられる。本研究の手法によると、観察場面とタスク再現場面において、同様の視覚機能が利用可能であるから、この計画法も学習・再現時に一般的に有用となるであろう。

人間の幼児が三歳時と五歳時に質問に答えられる道具利用法に関する調査結果 [120] によると、人間は、バケツに入れる、飲む、長靴をはくと言った、言わば物理的な道具利用を先に学習し、スポンジを洗う、ご飯をたく、花に水をやると言った、化学的な作用を含む道具利用、あるいは複数の手順を踏んだ道具利用を後に学習していくようである。ロボットの知能もこのように、道具の利用目的や使い方に関して、より複雑な知識を獲得していけるようにすべきだと考えられる。その際に課題となる、観察しにくい作用の推定法に関しては、本研究で提案した変形や上下・包含・突入関係の変化を含む作用の視覚推定法を用い、さらに音などの情報も統合して、より複雑な物理法則によって成り立つ作用のマルチモーダルな推定法となっていくべきだと考えている。

将来、センサの技術が発達し、計算機の性能が向上しても、遮蔽関係にある物体を直接透視できるような機器や認識手法は現れないだろうと思われる。しかし一方で、道具で対象を操作する多くの場合において、道具と対象が近接・接触して作用しあい、上下・包含・突入関係が重要となることも変わらないであろう。道具と対象の全てを見ることができない中で、近接し接触する、重なる、包含する、突入するような関係を推定していくためには、本研究で述べたような時空間遮蔽近傍情報を用いる手法が重要たり続けると考えている。

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 博士課程に在学中、情報システム工学研究室（稲葉・岡田研究室，JSK）にておこなった研究をまとめたものです。

稲葉雅幸先生には、学部四年次で研究室に配属されて以来6年間、あらゆる面でお世話になりました。研究者としてのあるべき姿勢から、研究の組み立て方、実験例に関する具体的な構想まで、多くを教えていただき、本論文の内容に関して舵取りをしていただきました。常に視野を広く持ち、より上のレベルから研究を見直し、真に重要な研究を進めていく先生からは、まだ学ばねばならないことが多くあるように思います。今後ともご指導いただけますよう、よろしく願いいたします。

岡田慧先生にも学部時代より、研究内容からコーディング方法まで広くご指導いただきました。先生が中心となり整備してくださっているJSKのソフトウェア・ハードウェア基盤がなければ、本論文での実験は不可能であったと思います。常に妥協を許さない先生の姿勢を、私も今後より一層実践できるように努力していきたいと思ひます。

中村仁彦先生、國吉康夫先生、原田達也先生にも、学部時代より講義や演習にて大変お世話になりました。お忙しい中、学位審査をお引き受けくださり、先生方の貴重なご意見をいただけたことで、論文を有意義なものにできただけでなく、今後の研究の発展方向を見据えることができました。深く感謝申し上げます。

学部四年次よりお世話になり、特に修士課程在籍中より直接ご指導くださったのは山崎公俊先生でした。研究の基本的な進め方から具体的な実験の手法、論文の書き方、学会での発表方法、後輩の指導方法から生活リズムにいたるまで、先生自身が自ら手本となって、研究者のあるべき姿というものを教えてくださいました。山崎さんのような研究者となれるよう、今後とも努力していきたいと思ひます。

吉海智晃先生にも学部四年次より、JSKでの研究室生活方法を中心にご指導いただきました。修士二年次で、先生方とともに米国 Willow Garage に滞在し実験させていただいた経験は、博士課程における海外での学会発表や滞在型研究において、大変有益なものとなりました。花井亮先生には、特に博士一年次の共同研究にてご指導いただき、画像処理やコーディングに関する知識を深めることができました。矢口裕明先生には、特に博士三年次の産学共同研究にて、大変お世話になりました。研究内容や実験環境に関するアドバイスのみならず、遠隔地での研究では生活そのも

のも力強く支えていただきました。垣内洋平先生には、先生の豊富な知識から、ソフトウェア基盤から画像処理まで、幅広く教えていただきました。特に複合ビジョンシステムのキャリブレーションにて大変お世話になりました。野沢峻一先生にも学部四年次より、HRP-2の環境整備から研究生活のアドバイスにいたるまで、最も学年の近い先輩として常に的確なアドバイスをいただきました。山口真奈美さんにも、物品購入や出張をいつもサポートしていただき、集中して自分の研究に専念することができました。多くの先生方のご協力無しには、本論文の執筆は不可能であったと思います。改めて深く感謝申し上げます。

共同研究では、トヨタ自動車の皆様にも大変お世話になりました。特に山本貴史さん、宗玄清宏さん、土永将慶さんには、実験環境の整備から論文や特許へのアドバイス、生活のサポートまで、広範囲でお世話になりました。深く感謝申し上げます。

JSKでは先生方のみならず、多くの先輩方や後輩の皆さんにも大変お世話になりました。小倉崇先輩、寺田耕志先輩には、特に博士課程三年次において、経験に基づく研究への心構えのアドバイスから、実験環境の整備まで、大変お世話になりました。植田亮平先輩には、コーディングのプロとして、実験環境の整備からコーディング方法まで、特にPDDLでの物体状態記述に関するアドバイスを多くいただきました。白山翔太先輩にも、ズームヘッドの基盤整備や実験方法のご指導をいただき、高視力ヘッドでの実験を実現することができました。小林一也君には、同学年で共に博士論文を書く仲間として多くのアドバイスをいただきました。西野友博君には、著者の修士二年次から博士一年次のHRP-2VZ八号機の実験や共同研究にて、実験環境の整備から特徴量記述の議論、コーディングに至るまで毎日お世話になりました。大矢良輔君にも、著者の博士一年次から博士二年次にかけて毎日、HRP-2による実験準備から研究内容へのアドバイスまで、大変お世話になりました。大矢君のテーマであったHRP-2七号機による着衣支援実験も、事例ベースの研究として大変刺激になりました。また博士三年次の遠隔地での研究においても、生活のアドバイスや力強い支援をいただき、大変感謝しております。三宅健司君にも、博士二年次から博士三年時にかけて、布袋を使った実験を中心に、大変お世話になりました。博士三年次には共同研究で東京を留守にすることが多く迷惑をかけましたが、三宅君がやりきった布の道具利用実験は、今後とも発展させていきたいと思っています。田中義丸君、藤井祥平君、古田悠貴君にも、有益なアドバイスをいただきました。忙

しいスケジュールの中、本当にありがとうございました。

東京大学での研究への道を開いてくださった、著者の出身校の多くの先生方にも、感謝申し上げます。特に徳島の佐々木晃先生、大平和哉先生、兵庫の中山孝蔵先生、和田孫博先生は卒業後も変わらずいつも気にかけてくださり、大変心強く感じております。

剣道や居合道での稽古を通して、著者の研究生活を力強く支え正してくださった灘校剣道部の田中康俊先生、夢想清明会の沖野規匡先生をはじめ多くの先生方、馬場祐樹先輩をはじめ夢想清明会や灘校剣道部、日本大学居合道部の多くの皆様にも感謝申し上げます。

最後に、21年間という長い学生生活を身近で支えてくれた両親と家族、タロジロ、ペンタたちにも感謝し、謝辞の締めくくりとさせていただきます。

多くの皆様に支えられて博士論文を執筆できましたこと、大変有り難く思うと同時に、とても誇りに思います。皆様、本当にありがとうございました。

平成 25 年 12 月 13 日 長濱 虎太郎

発表文献等

投稿論文

- [i] Kotaro Nagahama, Tomohiro Nishino, Kimitoshi Yamazaki, Kei Okada and Masayuki Inaba. “Attention Region Detection and Tracking for Observing a Person Manipulating Day-to-day Objects,” *International Journal of Mechatronics and Automation*, Vol. 3, No. 1, pp. 58–67, 2013.
- [ii] Yoshiaki Watanabe, Kotaro Nagahama, Kimitoshi Yamazaki, Kei Okada and Masayuki Inaba, “Cooking Behavior with Handling General Cooking Tools based on a System Integration for a Life-sized Humanoid Robot,” *PALADYN Journal of Behavioral Robotics*, Vol. 4, Issue 2, pp. 63–72, 2013.

国際会議論文（査読有り）

主著

- [i] Kotaro Nagahama, Kimitoshi Yamazaki, Kei Okada and Masayuki Inaba. “Manipulation of Multiple Objects in Close Proximity Based on Visual Hierarchical Relationships,” in *Proceedings of the 2013 IEEE International Conference on Robotics and Automation*, pp. 1295–1302, 2013.
- [ii] Kotaro Nagahama, Kimitoshi Yamazaki, Kei Okada and Masayuki Inaba. “Hierarchical Estimation of Multiple Objects from Proximity Relationships Arising from Tool Manipulation,” in *Proceedings of the 2012 IEEE-RAS International Conference on Humanoid Robots*, pp. 666–673, 2012.
- [iii] Kotaro Nagahama, Tomohiro Nishino, Mitsuharu Kojima, Kimitoshi Yamazaki, Kei Okada and Masayuki Inaba. “End Point Tracking for a Moving Object with Several Attention Regions by Composite Vision System,” in *Proceedings of the IEEE International Conference on Mechatronics and Automation*, pp. 590–596, 2011.

共著

- [i] Kimitoshi Yamazaki, Kotaro Nagahama and Masayuki Inaba. “A Method of State Recognition of Dressing Clothes Based on Dynamic State Matching,” in *Proceedings of the 2013 IEEE/SICE International Symposium on System Integration*, pp. 406–411, 2013.
- [ii] Kimitoshi Yamazaki, Tomohiro Nishino, Kotaro Nagahama, Kei Okada and Masayuki Inaba. “A Vision System for Daily Assistive Robots Using Character Information in Daily Environments,” in *Proceedings of the 2013 IEEE/SICE International Symposium on System Integration*, pp. 901–906, 2013 (Best paper finalist).
- [iii] Kimitoshi Yamazaki, Ryo Hanai, Hiroaki Yaguchi, Kotaro Nagahama, Katsuyoshi Yamagami and Masayuki Inaba. “Comparing Investigation of Images and 3D Mesh Model For Categorizing Electric Appliance Components,” in *Proceedings of the 2013 IAPR International Conference on Machine Vision Applications*, pp. 427–430, 2013.
- [iv] Kimitoshi Yamazaki, Kotaro Nagahama and Masayuki Inaba. “Daily Clothes Observation from Visible Surfaces Based on Wrinkle and Cloth-Overlap Detection,” in *Proceedings of the IAPR Conference on Machine Vision Applications*, pp. 275–278, 2011.
- [v] Kimitoshi Yamazaki, Yoshiaki Watanabe, Kotaro Nagahama, Kei Okada and Masayuki Inaba. “Recognition and Manipulation Integration of a Daily Assistive Robot Working on Kitchen Environment,” in *Proceedings of the IEEE International Conference on Robotics and Biomimetics*, pp. 196–201, 2010.
- [vi] Wilma Alice Bainbridge, Shunichi Nozawa, Ryohei Ueda, Yohei Kakiuchi, Kotaro Nagahama, Kei Okada and Masayuki Inaba. “Understanding Expectations of a Robot’s Identity Through Multi-User Interactions,” in *Proceedings*

of the HRI 2011 Workshop on Exepctations in intuitive human-robot interaction, 2011.

国内会議論文（査読有り）

- [i] 長濱 虎太郎, 山崎 公俊, 岡田 慧, 稲葉 雅幸. “人の道具操作観察における複数領域追跡による重畳・随伴関係推定と操作機能分類,” 第17回ロボティクスシンポジウム予稿集, pp. 534–540, 2012.
- [ii] 矢口 裕明, 村瀬 和都, 長濱 虎太郎, 宗玄 清宏, 小倉 崇, 山本 貴史, 稲葉 雅幸. “インタラクティブ家具モデルを用いた生活支援ロボット HSR の家具操作教示手法,” 第19回ロボティクスシンポジウム予稿集, to appear.

国内会議論文（査読無し）

主著

- [i] 長濱 虎太郎, 三宅 健司, 山崎 公俊, 矢口 裕明, 岡田 慧, 稲葉 雅幸. “遮蔽近傍時刻を用いた見えない作用推定に基づく道具利用タスク学習法,” 第14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講演概要集, pp. 775–780, 2013.
- [ii] 長濱 虎太郎, 宗玄 清宏, 土永 将慶, 山本 貴史, 矢口 裕明, 稲葉 雅幸. “生活支援ロボット HSR のための遮蔽情報推定に基づく片づけタスク視覚学習・計画法,” 第14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講演概要集, pp. 781–786, 2013.
- [iii] 長濱 虎太郎, 西野 友博, 山崎 公俊, 岡田 慧, 稲葉 雅幸. “重畳関係と随伴性を用いた作用推定に基づくヒューマノイドの道具操作,” 日本機械学会ロボティクス・メカトロニクス講演会'12 講演論文集, 2A1-F04, 2012.

- [iv] 長濱 虎太郎, 山崎公俊, 岡田慧, 稲葉雅幸. “人の道具操作観察における機能分類のための重畳・随伴関係推定,” 第12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講演概要集, pp. 1833–1836, 2011. (優秀講演賞 受賞)
- [v] 長濱 虎太郎, 西野 友博, 山崎 公俊, 岡田 慧, 稲葉 雅幸. “生活支援ロボットによる日常物体操作観測のための複数対象の同時追跡と重畳関係推定,” 日本機械学会ロボティクス・メカトロニクス講演会’11 講演論文集, 2A1-H07, 2011.
- [vi] 長濱 虎太郎, 西野 友博, 山崎 公俊, 岡田 慧, 稲葉 雅幸. “ヒューマノイドの低次視覚機能による人の道具操作行動観測を目的とした時空間対象追跡,” 第28回日本ロボット学会学術講演会講演論文集, 1A2-2, 2010.
- [vii] 長濱 虎太郎, 白山 翔太, 植木 竜佑, 小島 光晴, 岡田 慧, 稲葉 雅幸. “ズームステレオを有する等身大ヒューマノイドにおける人と人の操作対象への注視誘導,” 日本機械学会ロボティクス・メカトロニクス講演会’10 講演論文集, 2P1-D19, 2010.
- [viii] 長濱 虎太郎, 野沢 峻一, 植田 亮平, 岡田 慧, 稲葉 雅幸. “ヒューマノイドの生活支援行動タスク模倣における作用点認識と全身行動生成,” 第27回日本ロボット学会学術講演会講演論文集, 1S1-05, 2009.
- [ix] 長濱 虎太郎, 岡田 慧, 稲葉 雅幸. “等身大ヒューマノイドの調理行動における視覚動作統合による操作対象物認識,” 第9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講演概要集, pp.737–738, 2008.

共著

- [i] 三宅 健司, 長濱 虎太郎, 矢口 裕明, 岡田 慧, 稲葉 雅幸. “双腕ロボットにより布袋へ物を入れるための開口操作と視覚状態認識,” 第14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講演概要集, pp.817–820, 2013.
- [ii] 矢口 裕明, 村瀬 和都, 長濱 虎太郎, 宗玄 清宏, 山本 貴史, 稲葉 雅幸. “生活支援ロボット HSR の遠隔操縦のための形状仮説に基づくインタラクティブ家具モデル構築手法,” 第31回日本ロボット学会学術講演会予稿集, 3C1-07, 2013.

- [iii] 大矢 良輔, 長濱 虎太郎, 三宅 健司, 山崎 公俊, 岡田 慧, 稲葉 雅幸. “衣類引っ掛かりの視覚検出による補正動作を用いたヒューマノイドロボットのズボン着衣支援タスクの実現,” 日本機械学会ロボティクス・メカトロニクス講演会'13 講演論文集, 2A2-Q13, 2013.
- [iv] 大矢 良輔, 長濱 虎太郎, 山崎 公俊, 岡田 慧, 稲葉 雅幸. “ヒューマノイドロボットによる着衣動作中の失敗状態を考慮した着衣支援行動の実現,” 第 30 回日本ロボット学会学術講演会予稿集, 3C2-8, 2012.
- [v] 山崎 公俊, 花井 亮, 矢口 裕明, 長濱 虎太郎, 山上勝義, 稲葉雅幸. “家電物品のカテゴリ識別問題における三次元形状と多視点画像の有用性検討,” 第 15 回画像の認識・理解シンポジウム予稿集, IS1-61, 2012.
- [vi] 西野 友博, 山崎 公俊, 長濱 虎太郎, 稲葉 雅幸. “注視領域可変カメラシステムを搭載したロボットによる環境中の文字列発見と読み取り,” 日本機械学会ロボティクス・メカトロニクス講演会'12 講演論文集, 2A1-F03, 2012.
- [vii] 大矢 良輔, 長濱 虎太郎, 山崎 公俊, 稲葉 雅幸. “状態遷移グラフを用いた衣類着脱支援のための布状態推定と着脱モデルの構築,” 情報処理学会第 74 回全国大会講演論文集, 4Q-4, 2012 (計測自動制御学会 学術奨励賞 受賞).
- [viii] 大矢 良輔, 長濱 虎太郎, 山崎 公俊, 稲葉 雅幸. “衣類着脱支援のための布内部領域の動的変化を用いた状態マッチング,” 第 12 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講演概要集, pp. 1728–1731, 2011.
- [ix] 吉海 智晃, 長濱 虎太郎, 小堀 浩子, 秋元 貴博, Isaac Anthony Alba, 伊藤 司, 岡田 慧, 稲葉 雅幸. “EusLisp をコアにした RTM と ROS の相互接続によるマルチロボットアプリケーションの実現,” 第 11 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講演概要集, pp. 1742–1745, 2010.
- [x] 白山 翔太, 長濱 虎太郎, 垣内 洋平, 小島 光晴, 岡田 慧, 稲葉 雅幸. “輻輳機構を有するズームカメラステレオビジョンにおけるキャリブレーションとズーム立体視の実現,” 日本機械学会ロボティクス・メカトロニクス講演会'10 講演論文集, 2A2-F15, 2010.

- [xi] 植木 竜佑, 長濱 虎太郎, 小島 光晴, 吉海 智晃, 岡田 慧, 稲葉 雅幸. “広視野高視力視覚を備えたヒューマノイドによる人との共同注意に基づく生活支援行動の実現,” 日本機械学会ロボティクス・メカトロニクス講演会’10 講演論文集, 2A2-B17, 2010.
- [xii] 渡辺 義明, 長濱 虎太郎, 山崎 公俊, 岡田 慧, 稲葉 雅幸. “力覚状態認識に基づく包丁を用いた野菜切断行動のヒューマノイドによる実現,” 日本機械学会ロボティクス・メカトロニクス講演会’10 講演論文集, 2A2-B18, 2010.
- [xiii] 渡辺 義明, 長濱 虎太郎, 山崎 公俊, 岡田 慧, 稲葉 雅幸. “表皮視覚認識と凹凸形状へのなじみ動作に基づく野菜皮む操作のヒューマノイドによる実現,” 日本機械学会ロボティクス・メカトロニクス講演会’10 講演論文集, 2A2-A16, 2010.
- [xiv] 渡辺 義明, 長濱 虎太郎, 山崎 公俊, 岡田 慧, 稲葉 雅幸. “一般調理道具を等身大ヒューマノイドの認識行動システム統合とサラダ調理の実現,” 日本機械学会ロボティクス・メカトロニクス講演会’10 講演論文集, 2A2-A17, 2010.

学位論文

- [i] 長濱 虎太郎. “日常生活支援ヒューマノイドにおける道具利用法獲得の為の注視実現法,” 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 修士論文, 2011.
- [ii] 長濱 虎太郎. “ヒューマノイドの道具操作における行動獲得に関する研究,” 東京大学 工学部 機械情報工学科 卒業論文, 2009.

その他講演

- [i] 長濱 虎太郎. “生活支援ヒューマノイドの道具利用法学習における視覚処理手法,” レクチャ, 第1回 R2P/IST リーディング・リサーチャー・カフェ, 東京大学本郷キャンパス, 2012.6.25.

- [ii] 長濱 虎太郎. “物体操作の行動観察における複数対象の同時追跡予測と重畳関係推定,” 話題提供, 2010 年度第 3 回山彦シンポジウム, 国立オリンピック記念青少年総合センター, 2011.3.5.

発明・特許

- [i] 長濱 虎太郎 他. 出願番号: 特 2013-257666, 「物体認識装置、ロボット及び物体認識方法」, 2013 年 12 月 13 日出願.

受賞

- [i] 長濱 虎太郎, 三宅 健司, 山崎 公俊, 矢口 裕明, 岡田 慧, 稲葉 雅幸. “遮蔽近傍時刻を用いた見えない作用推定に基づく道具利用タスク学習法,” 第 14 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 SI2013 優秀講演賞, 2013.12.
- [ii] Kimitoshi Yamazaki, Tomohiro Nishino, Kotaro Nagahama, Kei Okada and Masayuki Inaba. “A Vision System for Daily Assistive Robots Using Character Information in Daily Environments,” Best paper finalist at 2013 IEEE/SICE International Symposium on System Integration, 2013.12.
- [iii] 長濱 虎太郎, 山崎 公俊, 岡田 慧, 稲葉 雅幸. “人の道具操作観察における機能分類のための重畳・随伴関係推定,” 第 12 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 SI2011 優秀講演賞, 2011.12.
- [iv] 長濱 虎太郎. 日本機械学会畠山賞, 日本機械学会, 2009.3.24.
- [v] 長濱 虎太郎. 東京大学 平成 20 年度工学部長賞, 東京大学工学部, 2009.3.24.

参考文献

- [1] A. Amagai and K. Takase. “Implementation of dynamic manipulation with visual feedback and its application to pick and place task”. In *Proceedings of the IEEE International Symposium on Assembly and Task Planning*, pp. 344–350, 2001.
- [2] Noriaki Ando, Takashi Suehiro, Kosei Kitagaki, Tetsuo Kotoku, and Woo-Keun 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.
- [3] T. Asfour, K. Regenstein, P. Azad, J. Schröde, , and R. Dillmann. “ARMAR-III: A HUMANOID PLATFORM FOR PERCEPTION-ACTION INTEGRATION”. In *Proceedings of the 2nd International Workshop on Human-Centered Robotic Systems*, 2006.
- [4] Tamim Asfour, Kai Welke, Pedram Azad, Ales Ude, and Rüdiger Dillmann. “The Karlsruhe Humanoid Head”. In *Proceedings of the 8th IEEE-RAS International Conference on Humanoid Robots*, pp. 447–453, 2008.
- [5] Eren Erdal Askoy, Alexey Abramov, Florentin Wörgötter, and Babette Dellen. “Categorizing object-action relations from semantic scene graphs”. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 398–405, 2010.
- [6] ASUSTeK Computer Inc. “ASUS Xtion PRO LIVE webpage”, 2013 閲覧. http://www.asus.com/Multimedia/Xtion_PRO_LIVE/.
- [7] Haseru Azuma, Yohei Kakiuchi, Manabu Saito, Kei Okada, and Masayuki Inaba. “View-base Multi-touch Gesture Interface for Furniture Manipulation Robots”. In *Proceedings of the IEEE Workshop on Advanced Robotics and Its Social Impacts*, 2011.

- [8] Baumer. “Baumer industrial cameras’ webpage”, 2013 閲覧. <http://www.baumer.com/de-en/products/identification-image-processing/industrial-cameras/tx-series/tx-series-poe/>.
- [9] Michael Beetz, Ulrich Klank, Ingo Kresse, Alexis Maldonado, Lorenz Mösenlechner, Dejan Pangercic, Thomas Rühr, and Moritz Tenorth. “Robotic Roommates Making Pancakes”. In *Proceedings of the 11th IEEE-RAS International Conference on Humanoid Robots*, pp. 529–536, 2011.
- [10] Aude Billard, Sylvain Calinon, Ruediger Dillmann, and Stefan Schaal. *Survey: Robot Programming by Demonstration*, chapter 59, pp. 1–15. Handbook of Robotics, MIT Press, 2007.
- [11] Yuri Boykov and Vladimir Kolmogorov. “An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, pp. 1124–1137, 2004.
- [12] Gary Bradski and Adrian Kaehler. 詳解 OpenCV コンピュータビジョンライブラリを使った画像処理・認識. オライリージャパン, 2009.
- [13] Duane C. Brown. “Decentering Distortion of Lenses”. *PHOTOGRAMMETRIC ENGINEERING*, Vol. XXXII, No. 3, pp. 444–462, 1966.
- [14] Duane C. Brown. “Close-range camera calibration”. *PHOTOGRAMMETRIC ENGINEERING*, Vol. 37, No. 8, pp. 855–866, 1971.
- [15] Sylvain Calinon and Aude Billard. “Recognition and Reproduction of Gestures using a Probabilistic Framework combining PCA, ICA and HMM”. In *Proceedings of the 22nd International Conference on Machine Learning*, pp. 105–112, 2005.
- [16] Sylvain Calinon, Florent Guenter, and Aude Billard. “Goal-Directed Imitation in a Humanoid Robot”. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 300–305, 2005.

- [17] Sylvain Calinon, Florent Guenter, and Aude Billard. “On Learning, Representing and Generalizing a Task in a Humanoid Robot”. *IEEE Transactions on Systems, Man and Cybernetics, Part B, Special issue on robot learning by observation, demonstration and imitation*, Vol. 37, No. 2, pp. 286–298, 2007.
- [18] John Canny. “A Computational Approach to Edge Detection”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, pp. 679–698, 1986.
- [19] Tan Fung Chan and Rajiv V. Dubey. “A weighted least-norm solution based scheme for avoiding joint limits for redundant joint manipulators”. In *Robotics and Automation, IEEE Transactions on*, pp. 286–292, 1995.
- [20] Jonathan H. Connell. “Generating and Generalizing Models of Visual Objects”. *Artificial Intelligence*, Vol. 31, No. 2, pp. 159–183, 1987.
- [21] TOYOTA MOTOR CORPORATION. “Partner Robot — TOYOTA MOTOR CORPORATION GLOBAL WEBSITE”, 2013 閱覽. http://www.toyota-global.com/innovation/partner_robot/.
- [22] Rosen Diankov. “Automated Construction of Robotic Manipulation Programs”. PhD thesis, Carnegie Mellon University, Robotics Institute, 2010.
- [23] Christian Dornhege and Andreas Hertle. “Integrated Symbolic Planning in the Tidyup-Robot Project”. In *Proceedings of the AAAI Spring Symposium - Designing Intelligent Robots: Reintegrating AI II*, 2013. <http://www.informatik.uni-freiburg.de/~ki/papers/dornhege-hertle-aaai13ss.pdf>.
- [24] Zoran Duric, Jeffrey A. Fayman, and Ehud Rivlin. “Function from motion”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 6, pp. 579–591, 1996.

- [25] Richard E. Fikes and Nils J. Nilsson. “STRIPS: A NEW APPROACH TO THE APPLICATION OF THEOREM PROVING TO PROBLEM SOLVING”. In *Proceedings of the 1971 International Joint Conference on Artificial Intelligence*, pp. 608 – 620, 1971.
- [26] Willow Garage. “Overview — Willow Garage”, 2013 閱覽. <https://www.willowgarage.com/pages/pr2/overview>.
- [27] D. M. Gavrila. “Multi-feature Hierarchical Template matching Using Distance Transforms”. In *Proceedings of the 14th International Conference on Pattern Recognition*, pp. 439–444, 1998.
- [28] Malik Ghallab, Adele Howe, Craig Knoblock, Drew McDermott, Ashwin Ram, Manuela Veloso, Daniel Weld, and David Wilkins. “PDDL – The Planning Domain Definition Language”. Technical Report CVC TR-98-003/DCS TR-1165, Yale Center for Computational Vision and Control, 1998.
- [29] Michael Gienger, Herbert Jansen, and Christian Goeric. “Exploiting Task Intervals for Whole Body Robot Control”. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2484–2490, 2006.
- [30] Helmut Grabner, Jiri Matas, Luc Van Gool, and Philippe Cattin. “Tracking the Invisible: Learning Where the Object Might be”. In *Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1285–1292, 2010.
- [31] Stefan Hinterstoisser, Stefan Holzer, Cedric Cagniart, Slobodan Ilic, Kurt Konolige, Nassir Navab, and Vincent Lepetit. “Multimodal Templates for Real-Time Detection of Texture-less Objects in Heavily Cluttered Scenes”. In *Proceedings of the 2011 IEEE International Conference on Computer Vision*, pp. 858–865, 2011.
- [32] Hirohisa Hirukawa, Fumio Kanehiro, Kenji Kaneko, Shuuji Kajita, Kiyoshi Fujiwara, Yoshihiro Kawai, Fumiaki Tomita, Shigeoki Hirai, Kazuo Tanie,

- Takakatsu Isozumi, Kazuhiko Akachi, Toshikazu Kawasaki, Shigehiko Ota, Kazuhiko Yokoyama, Hiroyuki Handa, Yutaro Fukase, Jun ichiro Maeda, Yoshihiko Nakamura, Susumu Tachi, and Hirochika Inoue. “Humanoid robotics platforms developed in HRP”. In *Proceedings of the 2003 IEEE-RAS International Conference on Humanoid Robots*, pp. 165–175, 2003.
- [33] Honda Motor Co., Ltd. “Honda — ASIMO”, 2013 閲覧. <http://www.honda.co.jp/ASIMO/>.
- [34] Katsushi Ikeuchi and Takashi Suchiro. “Towards an assembly plan from observation. I. Assembly task recognition using face-contact relations (polyhedral objects)”. In *Proceedings of the 1992 IEEE International Conference on Robotics and Automation*, Vol. 3, pp. 2171–2177, 1992.
- [35] Tetsunari Inamura, Kei Okada, Satoru Tokutsu, Naotaka Hatao, Masayuki Inaba, and Hirochika Inoue. “HRP-2W: A humanoid platform for research on support behavior in daily life environments”. *Robotics and Autonomous Systems*, Vol. 57, No. 2, pp. 145–154, 2009.
- [36] iRobot Corporation. “iRobot Roomba Vacuum Cleaning Robot”, 2013 閲覧. <http://www.irobot.com/us/learn/home/roomba.aspx>.
- [37] Ken Ito and Shigeyuki Sakane. “Robust View-based Visual Tracking with Detection of Occlusions”. In *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, pp. 1207–1213, 2001.
- [38] Hiroyasu Iwata and Shigeki Sugano. “Design of human symbiotic robot TWENDY-ONE”. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 580–586, 2009.
- [39] Hiroyasu Iwata and Shigeki Sugano. “Design of Human Symbiotic Robot TWENDY-ONE”. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 580–586, 2009.

- [40] R. A. Jarvis. “A Perspective on Range Finding Techniques for Computer Vision”. *IEEE Transactions Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, No. 2, pp. 122–139, 1983.
- [41] Kenji Kaneko, Fumio Kanehiro, Shuuji Kajita, Hirohisa Hirukawa, Toshikazu Kawasaki, Masaru Hirata, Kazuhiko Akachi, and Takakatsu Isozumi. “Humanoid Robot HRP-2”. In *Proceedings of 2004 IEEE-RAS International Conference on Robotics and Automation*, pp. 1083–1090, 2004.
- [42] Emil Keyder and Héctor Geffner. “The FF(ha) Planner for Planning with Action Costs”. In *Document for Sixth International Planning Competition (IPC 2008)*, 2008. <http://ipc.informatik.uni-freiburg.de/Planners>.
- [43] Nate Kohl and Petter Stone. “Policy Gradient Reinforcement Learning for Fast Quadrupedal Locomotion”. In *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, Vol. 3, pp. 2619–2624, 2006.
- [44] Yasuo Kuniyoshi, Masayuki Inaba, and Hirochika Inoue. “Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance”. *IEEE Transactions on Robotics and Automation*, Vol. 10, No. 6, pp. 799–822, 1994.
- [45] Yasuo Kuniyoshi, Nobuyuki Kita, Kazuhide Sugimoto, Shin Nakamura, and Takashi Suehiro. “A Foveated Wide Angle Lens for Active Vision”. In *Proceedings of the 1995 IEEE International Conference on Robotics and Automations*, pp. 2982–2988, 1995.
- [46] David G. Lowe. “Object Recognition from Local Scale-Invariant Features”. In *Proceedings of the 1999 IEEE International Conference on Computer Vision*, pp. 1150–1157, 1999.
- [47] Kazuhiro Maki, Noriaki Katayama, Nobutaka Shimada, and Yoshiaki Shirai. “Image-Based Automatic Detection of Indoor Scene Events and Interac-

- tive Inquiry”. In *Proceeding of the 19th International Conference on Pattern Recognition*, pp. 1–4, 2008.
- [48] Giorgio Metta, Giulio Sandini, David Vernon, Lorenzo Natale, and Francesco Nori. “The iCub humanoid robot: an open platform for research in embodied cognition”. In *Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems*, pp. 50–56, 2008.
- [49] Microsoft. “Kinect - Xbox.com”, 2013 閲覧. <http://www.xbox.com/en-us/kinect>.
- [50] Hiroyuki Miyamoto, Stefan Schaal, Francesca Gandolfo, Hiroaki Gomi, Yasuharu Koike, Rieko Osu, Eri Nakano, Yasuhiro Wada, and Mitsuo Kawato. “A Kendama Learning Robot Based on Bi-directional Theory”. *Neural Networks*, Vol. 9, No. 8, pp. 1281–1302, 1996.
- [51] Kotaro Nagahama, Tomohiro Nishino, Mitsuharu Kojima, Kimitoshi Yamazaki, Kei Okada, and Masayuki Inaba. “End Point Tracking for a Moving Object with Several Attention Regions by Composite Vision System”. In *Proceedings of IEEE International Conference on Mechatronics and Automation*, pp. 590–596, 2011.
- [52] Kotaro Nagahama, Tomohiro Nishino, Kimitoshi Yamazaki, Kei Okada, and Masayuki Inaba. “Attention Region Detection and Tracking for Observing a Person Manipulating Day-to-day Objects”. *International Journal of Mechatronics and Automation*, Vol. 3, No. 1, pp. 58–67, 2013.
- [53] Kotaro Nagahama, Kimitoshi Yamazaki, Kei Okada, and Masayuki Inaba. “Hierarchical Estimation of Multiple Objects from Proximity Relationships Arising from Tool Manipulation”. In *Proceedings of the 2012 IEEE-RAS International Conference on Humanoid Robots*, pp. 666–673, 2012.
- [54] Kotaro Nagahama, Kimitoshi Yamazaki, Kei Okada, and Masayuki Inaba. Manipulation of multiple objects in close proximity based on visual hierarchical

- relationships. In *Proceedings of the 2013 IEEE International Conference on Robotics and Automation*, pp. 1295–1302, 2013.
- [55] Yoshihiko Nakamura and Hideo Hanafusa. “Inverse Kinematic Solutions With Singularity Robustness for Robot Manipulator Control”. *Journal of Dynamic Systems, Measurement, and Control*, Vol. 108, No. 3, pp. 163–171, 1986.
- [56] Kei Okada, Mitsuharu Kojima, Satoru Tokutsu, Yuto Mori, Toshiaki Maki, , and Masayuki Inaba. “Task Guided Attention Control and Visual Verification in Tea Serving by the Daily Assistive Humanoid HRP2JSK”. In *Proceedings of the 2008 IEEE International Conference on Intelligent Robots and Systems*, pp. 1551–1557, 2008.
- [57] Kei Okada, Satoru Tokutsu, Takashi Ogura, Mitsuharu Kojima, Yuto Mori, Toshiaki Maki, and Masayuki Inaba. “Scenario controller for daily assistive humanoid using visual verification”. In *Proceedings of the 10th International Conference on Intelligent Autonomous Systems*, pp. 398–405, 2008.
- [58] Eiichi Ono, Hidehiko Okabe, Hisao Ichijo, Noboru Aisaka, and Hitoshi Akami. “ROBOT HAND WITH SENSOR FOR HANDLING CLOTH”. In *Proceedings of the IEEE International Workshop on Intelligent Robots and Systems*, pp. 995–1000, 1990.
- [59] Guillem Palou and Philippe Salembier. “OCCLUSION-BASED DEPTH ORDERING ON MONOCULAR IMAGES WITH BINARY PARTITION TREE”. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1093–1096, 2011.
- [60] D. Rosenberg, M. D. Levine, and S. W. Zucker. “Computing relative depth relationships from occlusion cues”. In *Proceedings of the 4th International Joint Conference on Pattern Recognition*, pp. 765–769, 1978.
- [61] Ros.org. “euslisp - ROS Wiki”, 2013 閲覧. <http://wiki.ros.org/euslisp>.

- [62] Ros.org. “ROS.org”, 2013 閲覧. <http://www.ros.org/>.
- [63] Mark Ross. “MODEL-FREE, STATISTICAL DETECTION AND TRACKING OF MOVING OBJECTS”. In *Proceedings of the 2006 IEEE International Conference on Image Processing*, pp. 557–560, 2006.
- [64] Hisashi Sugiura, Michael Gienger, Herbert Janssen, and Christian Goerick. “Real-Time Self Collision Avoidance for Humanoids by means of Nullspace Criteria and Task Intervals”. In *Proceedings of the 2006 IEEE-RAS International Conference on Humanoid Robots*, pp. 575–580, 2006.
- [65] Hisashi Sugiura, Michael Gienger, Herbert Janssen, and Christian Goerick. “Real-time collision avoidance with whole body motion control for humanoid robots”. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2053–2068, 2007.
- [66] Yuta Sugiura, Takeo Igarashi, Hiroki Takahashi, Tabare A. Gowon, Charith Lasantha Fernando, Maki Sugimoto, and Masahiko Inami. “Foldy: Graphical Instruction for A Garment Folding Robot”. In *Proceedings of the ACM SIGGRAPH 2009*, 2009.
- [67] Jun Takamatsu, Hiroshi Kimura, and Katsushi Ikeuchi. “Calculating Optimal Trajectories from Contact Transitions”. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 2, pp. 1547–1552, 2002.
- [68] Jun Takamatsu, Hirohisa Tominaga, Koichi Ogawara, Hiroshi Kimura, and Katsushi Ikeuchi. “Extracting Manipulation Skills from Observation”. In *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 1, pp. 584–589, 2000.
- [69] T. Tamei, T. Matsubara, A. Rai, and T. Shibata. “Reinforcement Learning of Clothing Assistance with a Dual-arm Robot”. In *Proceedings of the 11th IEEE-RAS International Conference on Humanoid Robots*, pp. 733–738, 2011.

- [70] Ryohei Ueda, Yohei Kakiuchi, Shunichi Nozawa, Kei Okada, and Masayuki Inaba. “Anytime Error Recovery by Integrating Local and Global Feedback with Monitoring Task States”. In *Proceedings of the 15th International Conference on Advanced Robotics*, p. TuII2.3, 2013.
- [71] Jur van den Berg, Stephen Miller, Ken Goldberg, and Pieter Abbeel. “Gravity-Based Robotic Cloth Folding”. In *Proceedings of the 9th International Workshop on the Algorithmic Foundations of Robotics*, pp. 1–16, 2010.
- [72] Peter von Kaenel, Chris Brown, and Dave Coombs. “Detecting Regions of Zero Disparity in Binocular Images”. Technical Report 388, The University of Rochester, 1991.
- [73] Kai Welke, Markus Przybylski, Tamim Asfour, and Rüdiger Dillmann. “Kinematic Calibration for Saccadic Eye Movements”. Technical report, Universität Karlsruhe, 2008.
- [74] Keenan A. Wyrobek, Eric H. Berger, H.F. Machiel Van der Loos, and J. Kenneth Salisbury. “Towards a Personal Robotics Development Platform: Rationale and Design of an Intrinsically Safe Personal Robot”. In *Proceedings of the 2008 IEEE International Conference on Robotics and Automation*, pp. 2165–2170, 2008.
- [75] Zhixing Xue, Steffen Ruehl, Andreas Hermann, Thilo Kerscher, and Ruediger Dillmann. “An Autonomous Ice-Cream Serving Robot”. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 3451–3452, 2011.
- [76] Kimitoshi Yamazaki, Kotaro Nagahama, and Masayuki Inaba. “Daily Clothes Observation from Visible Surfaces Based on Wrinkle and Cloth-Overlap Detection”. In *Proceedings of the 2011 IAPR Conference on Machine Vision Applications*, pp. 275–278, 2011.

- [77] Kimitoshi Yamazaki, Kotaro Nagahama, and Masayuki Inaba. “A Method of State Recognition of Dressing Clothes Based on Dynamic State Matching”. In *Proceedings of the 2013 IEEE/SICE International Symposium on System Integration*, pp. 406–411, 2013.
- [78] Kimitoshi Yamazaki, Tomohiro Nishino, Kotaro Nagahama, Kei Okada, and Masayuki Inaba. “A Vision System for Daily Assistive Robots Using Character Information in Daily Environments”. In *Proceedings of the 2013 IEEE/SICE International Symposium on System Integration*, pp. 901–906, 2013.
- [79] Kimitoshi Yamazaki, Ryohei Ueda, Shunichi Nozawa, Yuto Mori, Toshiaki Maki, Naotaka Hatao, Kei Okada, and Masayuki Inaba. “System Integration of a Daily Assistive Robot and its Application to Tidying and Cleaning Rooms”. In *Proceedings of The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1365–1371, 2010.
- [80] Kimitoshi Yamazaki, Ryohei Ueda, Shunichi Nozawa, Yuto Mori, Toshiaki Maki, Naotaka Hatao, Kei Okada, and Masayuki Inaba. “Tidying and Cleaning Rooms by a Daily Assistive Robot – An Integrated System for Doing Chores in Real World –”. *Journal of Behavioral Robotics*, Vol. 1, No. 4, pp. 231–239, 2011.
- [81] Kimitoshi Yamazaki, Yoshiaki Watanabe, Kotaro Nagahama, Kei Okada, and Masayuki Inaba. “Recognition and Manipulation Integration of a Daily Assistive Robot Working on Kitchen Environment”. In *Proceedings of the 2010 IEEE International Conference on Robotics and Biomimetics*, pp. 196–201, 2010.
- [82] Takashi Yoshimi, Naoyuki Iwata, Makoto Mizukawa, and Yoshinobu Ando. “Picking up Operation of Thin Objects by Robot Arm with Two-Fingered Parallel Soft Gripper”. In *Proceedings of the 2012 IEEE International Workshop on Advanced Robotics and its Social Impacts*, pp. 7–12, 2012.

- [83] Zhengyou Zhang. “A Flexible New Technique for Camera Calibration”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, pp. 1330–1334, 2000.
- [84] 杉本旭. “ロボットの教示と安全”. *日本ロボット学会誌*, Vol. 14, No. 3, p. 13, 1996.
- [85] 稲葉雅幸. “IRT 基盤創出プロジェクトにおけるロボットシステム研究”. 第 25 回日本ロボット学会学術講演会講演論文集, pp. 1L3–6, 2007.
- [86] 稲葉雅幸, 加賀美聡, 西脇光一. “岩波講座 ロボット学 ロボットアナトミー”. 岩波書店, 2005.
- [87] 藤田雅史, 小川原光一. “能動的な操作によって得られる連続画像を利用した特徴点に基づく衣類認識法”. 第 14 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講演概要集, pp. 1K1–6, 2013.
- [88] 河村雅人, 怡土順一, 竹村憲太郎, 松本吉央, 高松淳, 小笠原司. “視線方向知覚を考慮したアンドロイドの頭部・眼球動作”. 第 9 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予稿集, pp. 1019–1020, 2008.
- [89] 株式会社アフレル. Web Page: 教育版レゴ マインドストーム NXT, 2013 閲覧. <http://www.afrel.co.jp/lineup/nxt>.
- [90] 西野環. “ヒューマノイドにおける両眼視覚機構と感覚行動統合システムの実現”. 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 修士論文, 2008.
- [91] 西野環, 小島光晴, 吉海智晃, 岡田慧, 稲葉雅幸. “ズームレンズ搭載カメラを持つ輻輳可能なヒューマノイドヘッドの設計と実現”. *日本機械学会ロボティクス・メカトロニクス講演会'08 講演論文集*, pp. 2P1–F04, 2008.
- [92] 山本貴史, 齋藤史倫, 橋本国松, 池田幸一. “生活支援ロボット H S R の開発”. 第 30 回日本ロボット学会学術講演会予稿集, pp. 3C2–1, 2012.

- [93] 岡田慧, 野沢峻一, 植田亮平, 渡辺義明, 得津覚, 稲葉雅幸. “失敗からの復帰が可能な認識行動タスクの高次プランニングシステムにおける感覚に基づく動作の修正”. 第 27 回日本ロボット学会学術講演会予稿集, pp. 1L3-04, 2009.
- [94] 岡田慧, 小倉崇, 羽根田淳, 藤本純也, 佐川裕一, 市野利幸, 稲邑哲也, 稲葉雅幸. “空間行動プランナ機能に基づくヒューマノイドの高次模倣行動の観察と生成”. 第 23 回日本ロボット学会学術講演会予稿集, pp. 2H2-6, 2005.
- [95] 松本潔, 石川勝, 稲葉雅幸, 下山勲. “高齢社会におけるロボットによる生活支援技術, 東京大学 IRT 研究機構概要”, 2013 閲覧. <http://www.irt.i.u-tokyo.ac.jp/outline/index.shtml>.
- [96] 三宅健司. “双腕ロボットにおける柔軟な袋の視覚状態認識と開口物入れ操作に関する研究”. 東京大学大学院 情報理工学系研究科 創造情報学専攻 修士論文, 2014.
- [97] 三宅健司, 長濱虎太郎, 矢口裕明, 岡田慧, 稲葉雅幸. “双腕ロボットにより布袋へ物を入れるための開口操作と視覚状態認識”. 第 14 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講演概要集, pp. 817-820, 2013.
- [98] 金子健二, 金広文男, 森澤光晴, 三浦郁奈子, 中岡慎一郎, 原田研介, 梶田秀司. “サイバネティックヒューマン HRP-4C の開発”. 日本ロボット学会誌, Vol. 28, No. 7, pp. 853-864, 2010.
- [99] 長濱虎太郎. “日常生活支援ヒューマノイドにおける道具利用法獲得の為の注視実現法”. 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 修士論文, 2011.
- [100] 長濱虎太郎, 三宅健司, 山崎公俊, 矢口裕明, 岡田慧, 稲葉雅幸. “遮蔽近傍時刻を用いた見えない作用推定に基づく道具利用タスク学習法”. 第 14 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. 1K1-1, 2013.
- [101] 長濱虎太郎, 山崎公俊, 岡田慧, 稲葉雅幸. “人の道具操作観察における機能分類のための重畳・随伴関係推定”. 第 12 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. 3A2-4, 2011.

- [102] 長濱虎太郎, 山崎公俊, 岡田慧, 稲葉雅幸. “人の道具操作観察における複数領域追跡による重畳・随伴関係推定と操作機能分類”. 第17回ロボティクスシンポジウム予稿集, pp. 534-540, 2012.
- [103] 長濱虎太郎, 野沢峻一, 植田亮平, 岡田慧, 稲葉雅幸. “ヒューマノイドの生活支援行動タスク模倣における作用点認識と全身行動生成”. 第27回日本ロボット学会学術講演会講演論文集, pp. 1S1-05, 2009.
- [104] 長濱虎太郎, 宗玄清宏, 土永将慶, 山本貴史, 矢口裕明, 稲葉雅幸. “生活支援ロボット HSR のための遮蔽情報推定に基づく片づけタスク視覚学習・計画法”. 第14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. 1K1-2, 2013.
- [105] 長濱虎太郎, 西野友博, 山崎公俊, 岡田慧, 稲葉雅幸. “ヒューマノイドの低次視覚機能による人の道具操作行動観測を目的とした時空間対象追跡”. 第28回日本ロボット学会学術講演会講演論文集, pp. 1A2-2, 2010.
- [106] 長濱虎太郎, 西野友博, 山崎公俊, 岡田慧, 稲葉雅幸. “重畳関係と随伴性を用いた作用推定に基づくヒューマノイドの道具操作”. 日本機械学会ロボティクス・メカトロニクス講演会'12 講演論文集, pp. 2A1-F04, 2012.
- [107] 長濱虎太郎, 白山翔太, 植木竜佑, 小島光晴, 岡田慧, 稲葉雅幸. “ズームステレオを有する等身大ヒューマノイドにおける人と人の操作対象への注視誘導”. 日本機械学会ロボティクス・メカトロニクス講演会'10 講演論文集, pp. 2P1-D19, 2010.
- [108] 小川原光一, 高松淳, 木村浩, 池内克史. “観察に基づく手作業の獲得における視覚の利用”. 情報処理学会論文誌コンピュータビジョンとイメージメディア, Vol. 44, No. 17, pp. 13-23, 2003.
- [109] 小島光晴, 土永将慶, 山本貴史, 矢口裕明, 稲葉雅幸. “生活支援ロボット HSR のための家庭環境における LINE-MOD の適用とロバスト性向上”. 情報処理学会第75回全国大会講演論文集, pp. 3D-7, 2013.

- [110] 小島光晴, 白山翔太, 植木竜佑, 袖山慶直, 得津覚, 吉海智晃, 岡田慧, 稲葉雅幸. “広視野高視力を備えたアクティブステレオビジョンヘッドによる見守りシステム”. 第9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. 219–220, 2008.
- [111] 白石康星, 西田佳史, 本村陽一, 大川弥生, 溝口博. “国際生活機能分類を用いた日常生活プロトコルデータの正規化に基づく生活機能構造のモデル化と理解”. 電子情報通信学会技術研究報告. NC, ニューロコンピューティング, Vol. 109, No. 461, pp. 431–436, 2010.
- [112] 國吉康夫. “模倣の適応性と創発性：実世界における人間的知性の基盤”. 日本ロボット学会誌, Vol. 25, No. 5, pp. 671–677, 2007.
- [113] 國吉康夫, 井上博允, 稲葉雅幸. “人間が実演して見せる作業の実時間視覚認識とそのロボット教示への応用”. 日本ロボット学会誌, Vol. 9, No. 3, pp. 295–303, 1991.
- [114] 山崎公俊, 稲葉雅幸. “布地, しわ, 布の折れ重なりに着目した画像特徴量による無造作に置かれた布製品の個体識別”. 計測自動制御学会論文集, Vol. 49, No. 7, pp. 661–669, 2013.
- [115] 齋藤史倫, 橋本国松, 池田幸一, 山本貴史, 畠中規, 飯島浩. “生活支援ロボット HSR の試作と実証評価”. 第30回日本ロボット学会学術講演会予稿集, pp. 3C2–2, 2012.
- [116] 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.
- [117] 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”. 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 博士論文, 2013.
- [118] 高松淳, 小川原光一, 木村浩, 池内克史. “ロボットによる実行を目的とした人の組み立て作業の理解-接触状態遷移からの最適軌道の生成-”. 日本ロボット学会誌, Vol. 22, No. 6, pp. 752–763, 2004.

- [119] 今井順一, 金子正秀. “対象の自律的切替え機構を持つパーティクルフィルタによる遮蔽環境下での視覚追跡 (ニューラルネットワーク画像復元及び一般)”. 電子情報通信学会技術研究報告. NC, ニューロコンピューティング, Vol. 106, No. 500, pp. 31–36, 2007.
- [120] 及川尚子. “幼児の手の労働に関する発達心理学的研究”. 徳島大學學藝紀要 (教育科學), Vol. 16, pp. 61–84, 1968.
- [121] 中岡慎一郎, 中澤篤志, 金広文男, 金子健二, 森澤光晴, 比留川博久, 池内克史. “脚タスクモデルを用いた2足歩行ヒューマノイドロボットによる人の舞踊動作の再現”. 日本ロボット学会誌, Vol. 24, No. 3, pp. 388–399, 2006.
- [122] 小倉崇. “等身大ヒューマノイドにおける統合知能システムと家事支援行動の実現”. 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 博士論文, 2007.
- [123] 小倉崇, 岡田稲葉 雅幸. “注目点を持つ幾何モデルを利用したヒューマノイドの道具利用動作の生成法”. 第23回日本ロボット学会学術講演会予稿集, pp. 1F1–5, 2005.
- [124] 小倉崇, 神崎秀, 岡田稲葉 雅幸. “主体的に情報収集を行う日常生活支援ヒューマノイドの構成法”. 第25回日本ロボット学会学術講演会予稿集, pp. 3H2–1, 2007.
- [125] 川島貴美江, 山田美津子. “高齢者のデイサービスセンターにおける介護プログラムに関する一考察”. 静岡県立大学短期大学部研究紀要 18 - W号 (2004年度), Vol. 11, pp. 1–9, 2005.
- [126] 足立勝, 岸泰生, 亀井泉寿, 中村高幸, 横山和. “タウンマネジメントシステムを用いたロボットの開発”. 第8回システムインテグレーション部門講演会予稿集, pp. 939–940, 2007.
- [127] 長濱 虎太郎他. 物体認識装置、ロボット及び物体認識方法. 特許 出願番号: 2013-257666. 2013年12月13日出願.

- [128] 森大樹, 内海章, 大谷淳, 谷内田正彦, 中津良平. “非同期多視点画像による人物追跡システムの構築”. 電子情報通信学会論文誌, Vol. J84-DII, No. 1, pp. 102–110, 2001.
- [129] 大場光太郎, 大原賢一. “ユビキタス・ロボティクス”. 日本ロボット学会誌, Vol. 25, No. 4, pp. 505–508, 2007.
- [130] 和弘大塚, 直樹武川. “多視点観測に基づく複数物体の相互オクルージョン解析と逐次状態推定”. 情報処理学会研究報告. CVIM, [コンピュータビジョンとイメージメディア], Vol. 2003, No. 2, pp. 123–130, 2003.
- [131] 久原卓, 出口大輔, 高橋友和, 井手一郎, 村瀬洋. “調理映像中の繰返し調理動作の検出”. 電子情報通信学会論文誌, Vol. J94-D, No. 12, pp. 1983–1985, 2011.
- [132] 佐藤知正, 門脇啓太, 森武俊, 杉正夫, 福井類. “食器片付けロボットのための画像処理システム”. 第27回日本ロボット学会学術講演会予稿集, pp. 2R1–06, 2009.
- [133] 佐藤知正, 久保寺秀幸, 原田達也, 森武俊. “人間動作データの定性表現に基づく物体操作作業の分類”. 電子情報通信学会技術研究報告. NC, ニューロコンピューティング, Vol. 103, No. 392, pp. 67–72, 2003.
- [134] 独立行政法人産業技術総合研究所. “産総研：プレスリリース IC タグを用いた『知識分散型ロボット制御システム』を開発”, 2003. http://www.aist.go.jp/aist_j/press_release/pr2003/pr20031111_2/pr20031111_2.html.
- [135] 貞本敦史. “NEDO 生活支援ロボット実用化プロジェクトの概要”. 日本ロボット学会誌, Vol. 29, No. 9, pp. 750–764, 2011.
- [136] 鍋嶋厚太, 國吉康夫. “対象同定に基づく道具身体化の計算論”. 日本ロボット学会誌, Vol. 25, No. 6, pp. 897–905, 2007.
- [137] 石川博. “グラフカット (チュートリアル)”. 情報処理学会研究報告. CVIM, [コンピュータビジョンとイメージメディア], Vol. 2007, No. 31, pp. 193–204, 2007.

- [138] 井上博允. “ロボット言語の研究課題”. 日本ロボット学会誌, Vol. 2, No. 2, pp. 87–90, 1984.
- [139] 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.
- [140] 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.
- [141] 光永法明, スミスクリスチャン, 神田崇行, 石黒浩, 萩田紀博. “方策勾配型強化学習によるロボットの対人行動の個人適応”. 日本ロボット学会誌, Vol. 24, No. 7, pp. 820–829, 2006.
- [142] 西野友博. “注視領域可変機構を有するロボットにおける環境中の文字発見とその利用に関する研究”. 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 修士論文, 2012.
- [143] 西野友博, 山崎公俊, 長濱虎太郎, 稲葉雅幸. “注視領域可変カメラシステムを搭載したロボットによる環境中の文字列発見と読み取り”. 日本機械学会ロボティクス・メカトロニクス講演会'12 講演論文集, pp. 2A1–F03, 2012.
- [144] 武蔵裕子. “料理道具の便利帳”. 大泉書店, 2009.
- [145] 矢口裕明, 村瀬和都, 長濱虎太郎, 宗玄清宏, 山本貴史, 稲葉雅幸. “生活支援ロボット HSR の遠隔操縦のための形状仮説に基づくインタラクティブ家具モデル構築手法”. 第 31 回日本ロボット学会学術講演会予稿集, pp. 3C1–07, 2013.
- [146] 山口雄紀, 信田春満, 西出俊, 奥乃博, 尾形哲也. “神経回路モデルを用いた道具身体化モデルによる道具機能表現の獲得”. 情報処理学会第 74 回全国大会講演論文集, pp. 4Q–3, 2012.
- [147] 木下和樹, 油田信一. “食器片付けロボットのための画像処理システム”. 第 27 回日本ロボット学会学術講演会予稿集, pp. 2R1–06, 2009.

- [148] 白山翔太. “広視野高視力ズームステレオの注視制御に基づく日常生活支援ヒューマノイド行動の実現”. 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 修士論文, 2010.
- [149] 白山翔太, 長濱虎太郎, 垣内洋平, 小島光晴, 岡田慧, 稲葉雅幸. “輻輳機構を有するズームカメラステレオビジョンにおけるキャリブレーションとズーム立体視の実現”. 日本機械学会ロボティクス・メカトロニクス講演会'10 講演論文集, pp. 2A2-F15, 2010.
- [150] 白山翔太, 植木竜佑, 小島光晴, 得津覚, 吉海智晃, 岡田慧, 稲葉雅幸. “広視野・高視力ヒューマノイド hrp2-vz の開発と柔軟な注視制御の実現”. 第 27 回日本ロボット学会学術講演会講演論文集, pp. 1S3-08, 2009.
- [151] 白山翔太, 植木竜佑, 小島光晴, 得津覚, 吉海智晃, 岡田慧, 稲葉雅幸. “広視野・高視力ヒューマノイドによる対話的提示動作認識を用いた日用品管理”. 第 27 回日本ロボット学会学術講演会講演論文集, pp. 2S1-04, 2009.
- [152] 白山翔太, 植木竜佑, 小島光晴, 袖山慶直, 岡田慧, 稲葉雅幸. “見守りシステムのための広視野高視力アクティブステレオ視覚システムの実現”. 日本機械学会ロボティクス・メカトロニクス講演会'09 講演論文集, pp. 1P1-C14, 2009.
- [153] 白山翔太, 植木竜佑, 小島光晴, 袖山慶直, 岡田慧, 稲葉雅幸. “広視野高視力アクティブステレオ視覚システムを用いた物品登録と探索”. 日本機械学会ロボティクス・メカトロニクス講演会'09 講演論文集, pp. 1P1-C15, 2009.

以上

1p ~ 280p 完

博士論文

平成 25 年 12 月 13 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
長濱 虎太郎