

# 卒業論文

身体像に基づく感覚運動予測による  
自己と環境の隨伴性認知発達モデル

2012年2月3日 提出

指導教員 國吉 康夫 教授

東京大学 工学部 機械情報工学科 4年  
100281 芝原 俊樹



# 概要

乳児は環境にとのインタラクションを通じて、自己や環境の認知機構を発達させていく。乳児は周囲の環境を自己の運動に対する環境の変化に基づいて認識していく。乳児の周囲には単純な反応を示すもの、複雑な反応を示すもの、乳児の運動とは関係なく変化するものなど様々である。また、環境は様々に変化し、それとともに運動と環境の変化の関係性も変わっていく。このような環境で適応的に行動するためには状況の切り分けが重要となる。

この様な認知発達過程において、自己を対象とした運動から環境を対象とした運動への遷移が観察される。この運動発達は状況の切り分けを効率的に導いていると考える。しかし、この運動発達の原理は解明されていない。

また、乳児は周囲の環境に積極的に働きかける探索行動を頻繁におこなう。探索行動は行動によって報酬はもたらされず、明示的な動機づけが存在しない。探索行動の動機づけとして、内発的動機づけが考えられている。内発的モチベーションに基づく探索行動として、自己の身体が起点となり、その後自己の周囲の環境に探索の範囲が広がっていくことで自己を対象とした運動から環境を対象とした運動へ運動が遷移していくと考える。

感覚運動予測に基づく内発的モチベーションによる運動生成モデル化し、ロボットに実装してた。物体の存在する環境中でインタラクションさせ、運動の遷移と獲得された状況の切り分けを評価した。

運動性性が内発的モチベーションに基づく場合は、自己を対象とした運動から環境を対象とした運動への運動遷移が観察された。そして、状況の切り分けを導く可能性が示唆された。

内発的モチベーションは心理学で議論されている通り、単純なものから複雑なものへの運動遷移の要因となり、状況の切り分けにおける乳児の認知発達を導いていると考えられる。



# 目次

<b>第1章</b>	<b>序論</b>	<b>1</b>
1.1	はじめに . . . . .	1
1.2	目的 . . . . .	2
1.3	論文の構成 . . . . .	2
<b>第2章</b>	<b>乳児の運動発達と運動感覚予測</b>	<b>3</b>
2.1	乳児の発達 . . . . .	3
2.1.1	乳児の運動発達 . . . . .	4
2.2	乳児の認知発達 . . . . .	5
2.2.1	感覚運動の学習 . . . . .	5
2.2.2	感覚運動予測 . . . . .	7
2.3	学習モデルと脳内機構 . . . . .	9
2.3.1	強化学習 . . . . .	9
2.3.2	強化学習の脳内機構 . . . . .	10
2.3.3	競合学習 . . . . .	12
2.3.4	脳のモジュール性 . . . . .	13
2.3.5	exploitation and exploration dilemma . . . . .	15
2.4	乳児の探索行動と内発的動機づけ . . . . .	15
2.4.1	心理学における内発的動機づけ . . . . .	15
2.4.2	内発的モチベーションの脳内機構 . . . . .	16
2.4.3	内発的動機づけモデル . . . . .	16
2.4.4	馴化 . . . . .	17
2.5	認知発達に対するアプローチ . . . . .	18
<b>第3章</b>	<b>感覚運動予測に基づく運動生成モデル</b>	<b>21</b>
3.1	システムの構成 . . . . .	21
3.2	モジュール制御方式による競合学習 . . . . .	22

---

3.2.1	予測器の設計 . . . . .	22
3.2.2	運動の重みの更新 . . . . .	22
3.3	モジュールの統合と運動選択 . . . . .	24
<b>第 4 章 実験</b>		27
4.1	システム構成 . . . . .	27
4.2	実装 . . . . .	29
4.2.1	モジュールの競合学習 . . . . .	29
4.2.2	運動生成 . . . . .	29
4.2.3	感覚運動予測に用いる手法と特徴量 . . . . .	30
4.2.4	感覚運動予測に用いる特徴量 . . . . .	31
4.3	実験 . . . . .	32
4.4	運動遷移実験 . . . . .	32
4.4.1	運動遷移の評価 . . . . .	33
4.4.2	内発的動機づけモデル . . . . .	34
4.4.3	異なる動機づけモデルとの比較 . . . . .	41
4.4.4	運動遷移実験に関する考察 . . . . .	45
4.5	状況の切り分けに関する実験 . . . . .	47
4.5.1	状況の切り分けの評価方法 . . . . .	48
4.5.2	状況の切り分けの評価 . . . . .	49
4.5.3	状況の切り分けについての考察 . . . . .	51
<b>第 5 章 結論と展望</b>		53
5.1	結論 . . . . .	53
5.2	展望と課題 . . . . .	54
5.2.1	自己形成の段階 . . . . .	54
<b>謝辞</b>		57
<b>参考文献</b>		59

# 図目次

2.1	Pillow responses across two weeks by the three groups[58]. . . . .	6
2.2	実験の様子 [45]. . . . .	6
2.3	音のなる閾値の強さでおしゃぶりを吸った頻度 [45]. . . . .	6
2.4	input and output of PPC[9]. . . . .	7
2.5	Prediction feedback by an internal forward model[20] . . . . .	8
2.6	Change the body image by using a tool[6] . . . . .	9
2.7	ドーパミン細胞の報酬予測応答 [67] . . . . .	11
2.8	大脳基底核の TD 学習モデル [4] . . . . .	12
2.9	Multiple Paired Predictor-Controller architecture [13] . . . . .	13
2.10	道具使用の習得と小脳の活動 [11]. . . . .	14
2.11	異なる道具使用時の小脳の活動 [63] . . . . .	14
3.1	システム概要. . . . .	22
3.2	予測誤差に対する運動の重みの加算値. . . . .	23
3.3	獲得される運動の重み. . . . .	23
3.4	運動生成分布と運動生成. . . . .	25
3.5	馴化による運動生成分布の変化 . . . . .	25
3.6	1回の馴化による影響 . . . . .	25
3.7	システム構成. . . . .	26
4.1	layout joint of NAO[39] . . . . .	28
4.2	Camera position lateral . . . . .	29
4.3	Camera position top . . . . .	29
4.4	Length of links . . . . .	30
4.5	小さい予測誤差に対する動機づけ . . . . .	32
4.6	大きい予測誤差に対する動機づけ . . . . .	32
4.7	ボールを用いた実験の様子 . . . . .	33

4.8	運動の生起頻度 . . . . .	33
4.9	内発的動機づけモデルによる運動の生起頻度 . . . . .	35
4.10	内発的動機づけモデルによる予測誤差の時間変化 . . . . .	36
4.11	内発的動機づけモデルによる運動の重みが増加した回数の時間変化 . . . . .	36
4.12	内発的動機づけモデルによる各カテゴリの Z 値 . . . . .	36
4.13	学習初期状態での運動の生起頻度 . . . . .	37
4.14	学習初期状態での Z 値 . . . . .	38
4.15	予測誤差の時間変化 . . . . .	38
4.16	興味度が増加した運動の回数の時間変化 . . . . .	38
4.17	学習後期での各カテゴリの運動の生起頻度 . . . . .	40
4.18	学習後期での各カテゴリの Z 値 . . . . .	40
4.19	予測誤差の時間変化 . . . . .	41
4.20	興味度を増加させる運動の回数の時間変化 . . . . .	41
4.21	予測誤差の小さい運動に対する重みの増加モデル . . . . .	42
4.22	運動の生起頻度 . . . . .	43
4.23	予測誤差の時間変化 . . . . .	43
4.24	興味度が増加した回数の時間変化 . . . . .	43
4.25	小さい予測誤差に対する重みの増加モデルの Z 値 . . . . .	44
4.26	予測誤差の大きい運動の重み増加モデル . . . . .	44
4.27	運動の生起頻度 . . . . .	45
4.28	予測誤差の時間変化 . . . . .	45
4.29	重みが増加した回数の時間変化 . . . . .	45
4.30	予測誤差が大きい運動に対する重みの増加モデルにおける Z 値 . . . . .	46
4.31	ボールを動かす運動の生起頻度とボールを動かす運動の予測誤差 . . . . .	47
4.32	ボールを動かす運動の生起頻度と視野中に手を動かす運動の予測誤差 . . . . .	47
4.33	視野中に手を動かす運動の生起頻度とボールを動かす運動の予測誤差 . . . . .	47
4.34	視野中に手を動かす運動の生起頻度と視野中に手を動かす運動の予測誤差 . . . . .	47
4.35	ボールがない状態の実験の様子 . . . . .	48
4.36	ボールがある状態の実験の様子 . . . . .	48
4.37	提案モデル . . . . .	49
4.38	小さい予測誤差により興味度が増加するモデル . . . . .	50
4.39	大きい予測誤差により興味度が増加するモデル . . . . .	50
4.40	各モデルの乖離度 . . . . .	51
4.41	各モデルのばらつき . . . . .	51





# 表目次

2.1	発達段階 . . . . .	3
2.2	感覚運動期の各段階 ([66] を参考とした) . . . . .	18
4.1	Joint Range of NAO. . . . .	28
4.2	Used joint and joint angle. . . . .	29
4.3	各カテゴリの運動数. . . . .	34
4.4	運動の生起頻度と予測誤差の相関係数と p 値. . . . .	46



# 第1章

## 序論

### 1.1 はじめに

乳児は生まれてから環境とのインタラクションを通じて認知機能を発達させていく。乳児はお腹が空いたときに泣くことや生得的な反射が支配的な状態から、身体的な発達とともに環境中の物体に働きかけ、操作方法を獲得したり、他者と興味の物体を共有できるようになったり、約1年半の間に目覚ましい速度で発達する [61]。

環境の本質的な特性は環境への働きかけ方を規定し、人間はそれに応じて適応的に働きかけ方を変化させ、環境に適応していく。乳児は環境への表象的な認知が獲得されていないと考えられている。そのため、感覚刺激による認知が支配的であり、自己の運動によって得られる感覚刺激をもとに環境へ適応し、環境への理解を深めていくと考えられる。このように運動感覚間の関係性を学習することで、周囲の物体、自己、他者への理解が深まり、表象的知能へと発達していく足がかりとなる [64]。

感覚運動の関係性は、自分の体の運動と腕を見ているときの視覚刺激のように完全に一致し単純な関係性を示すものから、ボールを転がしたときのボールの動きのように少し複雑な関係のもの、他者とのインタラクションのように関係性が複雑なものまで多岐に渡る。さらに、乳児の周囲を取り巻く環境は状況によっておもちゃが置いてある場合や、養育者がいる場合など様々であり時間と共に変化していく。変化する環境中では状況を切り分け、区別した上で学習することが、認知構造を獲得する上で重要な役割を果たしていると考えられる。

乳児期の行動として、自分のゆびをしゃぶる行動、ガラガラを振る行動、ボールを転がす行動などを反復することをよく観察することができる。このような行動の中で乳児は自分の運動に対する感覚刺激を楽しんでいるように思える。運動に対する感覚刺激を探索する行動を反復的に行うことで学習が促進され、認知発達が加速されていると考える。

この認知発達過程において、乳児は自分の手をよく見るような自己を対象とした運動から、周囲の物体へ手を伸ばして遊ぶような自己以外の環境を対象とした運動へ変化していく。この運動発達は乳児の認知機能の発達において状況の切り分けを促し、学習を導いていると考える。しかし、この運動発達の原理は解明されていない。

乳児の運動発達において環境に積極的に働きかける探索的行動が、乳児の認知発達を適切に導く要因であると考える。探索行動の起点として自己の身体をまず学習し、その後より複雑な環境へ探索行動の範囲が広がっていくと推測し、その探索行動が状況の切り分けを適切に行いややすい環境を作り出していると考えることができる。

## 1.2 目的

本研究の目的は、感覚運動予測に基づく内発的動機づけによる運動生成をロボットに実装し、環境とインターラクションにおける運動遷移を観察し、その中に獲得される自己や環境の認知機構を評価する。そこから内発的動機づけによる乳児の運動発達への影響とそれによって導かれる認知機構の発達について考察する。

## 1.3 論文の構成

本論文は5章からなる。以下に各章の概要を示す。

第1章は本研究の概要と目的を述べた。

第2章は、乳児の運動発達、認知発達に関する知見を考察し、本研究の認知発達に対するアプローチを説明する。

第3章では、感覚運動の学習システムとそれに基づく運動生成モデルについて説明する。

第4章では、提案モデルの検証実験における実験セットアップと実験結果をまとめる。

最後に、第5章では、実験結果に対する考察をし、研究の課題・応用・展望をまとめる。

## 第2章

# 乳児の運動発達と運動感覚予測

本章では、乳児の運動発達と認知発達に関する心理学、脳科学の知見を考察する。最後に、本研究の認知発達に対するアプローチを述べる。

### 2.1 乳児の発達

人間の発達は連続的な過程であるが、ある特定の時期において飛躍的な変化が生じる。発達の時期には個人差があるが、その順序は固定的である。この変化のに基づいて発達を理解する指標として発達段階が考えられてきた。発達段階の区分は学者によって様々であるが、社会的な基準における一般的な区分を Table2.1 に示す [64]。乳児期は出生直後から歩行、初語の出現、離乳のほぼ完了などが観察される 1 歳半頃までと定義される。

乳児期では、約 1 年半の間に身体を思い通りに動かせない状態から、頭部から下肢に向かってコントロールできるようになり、歩行ができるようになる。また、運動発達とともに認知機能も発達していく、感覚運動的知能から物体の永続性を理解し表象

Table2.1 発達段階

発達段階	年齢
乳児期	0 歳～ 1 歳半頃まで
幼児期	1 歳半頃～ 6 歳頃まで
児童期	6 歳頃～ 12 歳頃まで
青年期	12 歳頃～ 20 歳頃まで
成人前期	20 歳頃～ 40 歳頃まで
成人中期	40 歳頃～ 65 歳頃まで
老年期	65 歳頃以降

的知能へ発達していく足がかりとなる。

### 2.1.1 乳児の運動発達

新生児の運動は、外界からの刺激に対して反射的に運動する原始反射と自発運動が中心的である [36]。

随意運動が始まる4ヶ月前まで、自発運動として外界からの刺激と無関係に、一定の時間間隔をあいて自発的に生じていると考えられる運動が見られ、四肢はランダムで可動域が広く、誇張された力強い動きをする [3]。自発運動は神経系の自発活動によって引き起こされていると考えられている。この運動は特定のパターンによって特徴づけることが難しいため、ジェネラルムーブメントと呼ばれる [37]。新生児期のジェネラルムーブメントは全身を大きく動かし、動きが予測できないライジングと呼ばれるものである。2ヶ月頃になると全身の屈伸をなめらかに繰り返すフィジエティーと呼ばれるパターンに変化する。ジェネラルムーブメントの計測も行われていて、生後1ヶ月児の運動はとても複雑であるが、2ヶ月児では単純で周期的な運動となる。3から4ヶ月児はまた複雑な運動を行うようになる [10]。

代表的な原始反射として、頭部が突然動かされたときに、上肢を伸ばして腕を開き、その後腕を曲げ、つかまるような動きをするモロー反射などがあげられる。他には仰向けの状態において、顔が向いている側の腕が伸展し、反対側の腕が屈曲する非対称性頸反射(ATNR)などがある。これらの反射は生後4~5ヶ月頃に消失する [3]。

乳児期の自発運動や原始反射は環境に適応した運動の学習や認知発達に関わっている。乳児の腕にモビールから伸びた紐を結び、腕の動きに随伴してモビールが動くようにしたときに、自発運動を手がかりに運動が獲得される。この獲得された運動が月齢ごとにどのように異なるか調べられている。2ヶ月児では体全体を動かすようになり、3ヶ月児では両腕を動かすようになり、4ヶ月児では紐がつながれた腕のみを動かすようになる。乳児に観察される運動は発達に従って、体全体を使った動きから、体の各部分を利用する動きへと発達していく [57]。原始反射であるATNRの状態において手が視野中で玩具に触れることがあり、ここから視覚と自己の身体との関係が始まること。

乳児に最初に見られる随意運動は2ヶ月頃の追視運動であるといわれている [65]。3ヶ月頃首がすわるようになり、そのことによって腕を顔の前にもっていき、手をよく見るハンドリガード呼ばれる運動をするようになる。

4ヶ月児になると玩具が胸部上に下げられると全身でリーチングしようとする。この段階では正確にリーチすることはできず、これを繰り返すことで運動のコントロールを習得していく。5ヶ月になるとリーチし、把持することができるようになる。把

持したあとは口に持って行ったり、振ったり、打ったりする。

9ヶ月児になると姿勢維持に必要とされなくなった肩や前腕の多様な運動によって、手の届く範囲外のものにもリーチングするようになる。また随意的につかんだ物を放すことができるようになる。

## 2.2 乳児の認知発達

乳児は環境と接触しながら自分の行動と環境からの刺激の関係性を学習することで、環境に対しての認知を発達させていく。この過程は環境への探索的行動により進行する。

乳児の感覚運動の認知に関する心理学の実験と、予測に関する脳内機構を説明する。

### 2.2.1 感覚運動の学習

乳児は環境と接触しながら、周囲の物体を運動と結びつけながら認識し、区別していく。

#### 随伴性刺激の認知

乳児は自分の運動に随伴する刺激を繰り返し知覚すると、その運動をする頻度が増加する。Watson による実験では、生後 8 週の乳児に 1 日 10 分間、頭を押し当てるべビーべット上のモビールが 1 秒間回転する枕を与え、乳児の枕に頭を押し当てる頻度を 2 週間に渡って観察した。また、頭を押し当てても動かないモビールのグループと、押し当てるのと関係なく一定時間において動くモビールのグループでも観察し比較した。この実験では枕と連動してモビールが動くグループのみ運動の頻度が増加した Fig.2.1。頭を押し当てたときにモビールが動く時と動かないときがある場合、モビールが動く時間が最初の実験の 40% または 60% の時に、運動の頻度の増加は見られなかった [58]。

#### 随伴性刺激の探索

新生児では自分の行動とそれに伴う刺激の関連付けを行っていないが、2ヶ月児になると自分の行動とそれに伴う刺激の関連付けをおこなっている。Rochat は、新生児と 2 ヶ月児にある強さ以上でおしゃぶりを吸うと、その吸った強さに応じた高さの音が出る条件 (analog condition), 閾値以上で吸うとランダムな高さの音が出る条件 (non-analog condition) の 2 つの条件で乳児の行動を観察した。この 2 つの条件における口で吸う強さの調節のばらつきを、この時乳児が自発的に操作するか、単なる隨

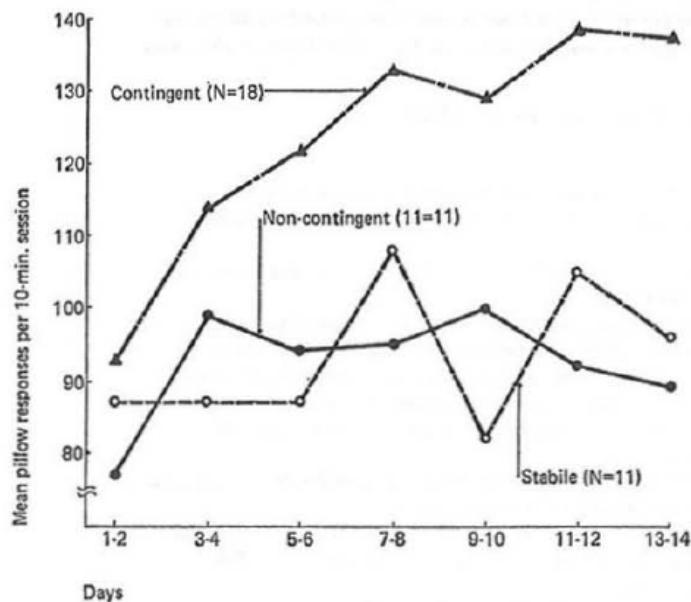


Fig.2.1 Pillow responses across two weeks by the three groups[58].

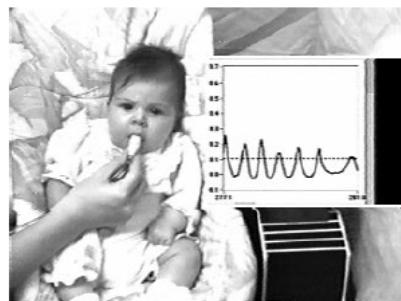


Fig.2.2 実験の様子 [45].

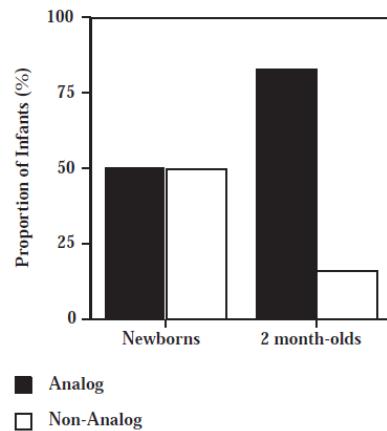


Fig.2.3 音のなる閾値の強さでおしゃぶりを吸った頻度 [45].

伴性の検出だけでなく、吸うことと音の因果関係を知覚しているかの指標とした。この実験の結果から、新生児では自分の行動と刺激の関係性を知覚していないが、2ヶ月児では自発的に操作し、自分の行動と刺激の関係性を知覚していると言える [45]。

#### 因果性の認知

乳児が因果関係が知覚できることを Leslie と Keeble は玉突き現象を用いて示した。赤い箱が緑の箱の方に動き、緑の箱に衝突し緑の箱がすぐに動き出す動画と、赤

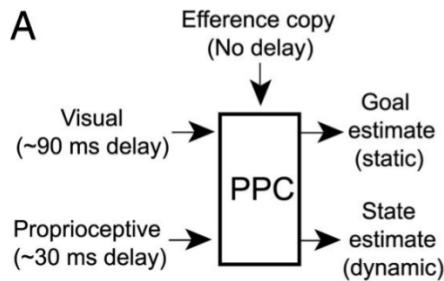


Fig.2.4 input and output of PPC[9].

い箱が緑の方に動き、緑の箱に衝突するが緑の箱は衝突後 0.5 秒後に動き始める動画を使って実験した。衝突後すぐに箱が動き出す動画では因果関係があると考えられるが、0.5 秒後に動き出す動画では因果関係が認められない。6 ヶ月児に一方の動画だけ見せて馴化させる。その後動画を逆再生して乳児に見せる。このとき、衝突後すぐに箱が動き出す動画の逆再生のみ新しい因果関係となる。衝突後すぐに動き出す動画の逆再生の方が 0.5 秒後に動き出す動画よりも脱馴化した。乳児は因果関係を知覚していると考えられる [32]。

### 2.2.2 感覚運動予測

人間は感覚運動ループに遅延があるにもかかわらず、素早く適切な運動の制御が可能である。この遅延に対処するために、脳内には次の状態を推定する順モデルを内部に持ち、使われていると考えられている。

運動する際には、脳から運動指令の信号が出されて筋肉に伝えられる。また、運動をスムーズに制御するために、運動指令によってどのような身体の状態になるか予測するために運動指令のコピーが使われる。これは遠心性コピーと呼ばれる [28]。

霊長類の脳において、後頭頂葉は感覚情報と運動情報の機能的インターフェイスとなっている。感覚情報の入力として、視覚と固有感覚の経路が知られている。運動情報としては、運動コマンドの信号の遠心性コピーが運動前野からほぼ時間遅延なく入力される。内部のフォワードモデルを用い現在及び直後の次の状態を見積もっていると提案されている [9]。

遠心性コピーはほぼ時間遅延なく後頭頂葉に入力されるが、他の感覚刺激は時間遅延があり、遅延も感覚器官によって様々である。そのため、後頭頂葉では入力される刺激の時間調整もされていると考えられている。

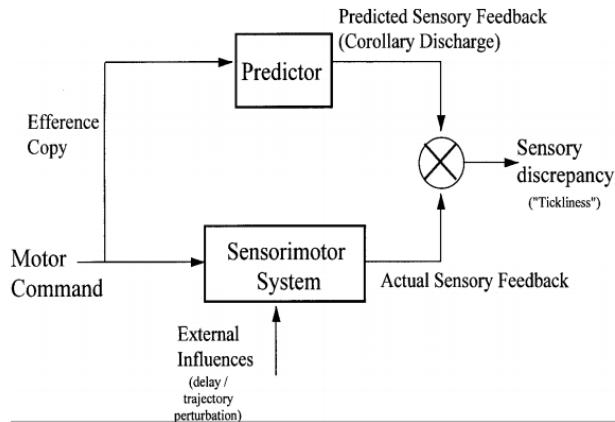


Fig.2.5 Prediction feedback by an internal forward model[20] .

### 自己による運動の感覚フィードバックの抑制

頭頂葉の順モデルは運動制御に用いられるだけでなく、感覚刺激の抑制にも用いられている。他人にくすぐられるとくすぐったいが、自分でくすぐったときにはくすぐったくないという現象がある。

Blakemore らは左手の動きに合わせて右手をくすぐる機械を行い実験を行った。左手の動きと一致してくすぐる場合、100, 200, 300 ミリ秒遅れてくすぐる場合、自動でくすぐる場合を比較した。くすぐったさの程度は時間遅れのない場合より時間遅れが大きくなるとともに大きくなり、300 ミリ秒の遅れで自動的にくすぐる場合とほぼ同じ強さになった [20]。

また、この現象を脳機能イメージによって解析したところ、第二次体性感覚野や小脳が関与していることが分かった [21]。

### 運動主体感と統合失調症

運動主体感は運動を引き起こしているのは自分であるという感覚である [27]。統合失調症の症状として自分のしている運動を他者にさせられたと感じる症状がある。この症状の原因として、自己による運動の感覚フィードバックが正常に行われないことが考えられている [22]。

### 道具使用と身体イメージ

自己の身体保持感や運動主体感から自己の身体イメージが形成されると考えられている。この身体イメージは固定的なものではなく、状況に応じて変化する。サルの頭頂葉の PEA という領域には手に対する体性感覚の刺激と手の周囲の視覚刺激に反応

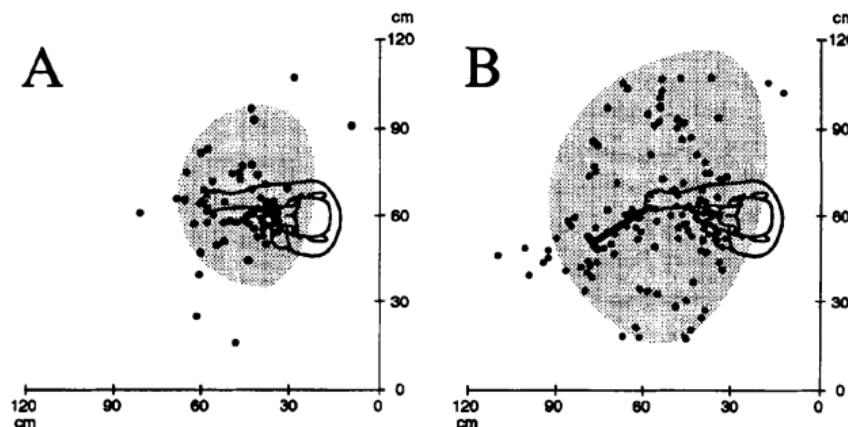


Fig.2.6 Change the body image by using a tool[6] .

するニューロンが存在する。このニューロンの活動をサルに熊手を使わせているときに記録した。ニューロンの活動は最初は手の周りの視覚刺激に対してだけであったが、熊手を使っていると熊手の先まで伸びることが明かになった [6]。

## 2.3 学習モデルと脳内機構

試行錯誤を通じ、周囲の環境を学習していくモデルとして強化学習、環境を分節化するモデルとして競合学習について解説する。そして、学習で重要な概念である exploitation と exploration のジレンマについてとりあげる。また、それについて脳との対応について考察する。

### 2.3.1 強化学習

強化学習とは、エージェントが試行錯誤を通じて環境に適応していく学習の枠組みである [54]。エージェントはある状態である行動を選択することで、他の状態に遷移する。このときそれに対応する報酬が与えられる。将来得る報酬を最大化することを目的として学習する。

今までの経験から状態の価値を推定し、それにしたがって最も価値の高い状態に遷移することを目指す。ある状態に対する価値を求める関数を価値関数とよぶ。ある状態でどの行動を選択するかをポリシーとよび、最適なポリシーを学習する。

環境のモデルを学習することで価値関数を求めるモデル同定強化学習と、直接価値関数を学習するモデルフリー強化学習がある。

### モデル同定強化学習

モデル同定強化学習は環境がマルコフ課程に従うことを仮定している。つまり、状態  $s$  で行動  $a$  を行った場合状態  $s'$  に遷移する確立は  $P(s'|s, a)$  となる。この時状態  $s$  の価値  $V(s)$  は(式2.1)となる。

$$V(s) = \max_a Q(s, a) \quad (2.1)$$

ここで、 $Q(s, a)$  は行動価値観数であり(式2.2)で定義される。

$$Q(s, a) = r(s, a) + \gamma \sum_{a'} P(s'|s, a)V(s') \quad (2.2)$$

$r(s, a)$  は報酬関数であり、 $\gamma$  は割引率である。状態遷移確率が既知の場合は動的計画法の問題である。

しかし、多くの場合状態遷移確率は未知である。そこで、環境のモデルを用いない強化学習がモデルフリー強化学習である。

### モデルフリー強化学習

モデルフリー強化学習の代表的な学習法として TD 学習がある[53]。TD 学習は報酬の予測誤差が 0 になるように価値関数を更新する手法である。予測誤差は TD 誤差とよばれ(式2.3)で表される。

$$\delta = r(s, a) + \gamma V(s') - V(s) \quad (2.3)$$

TD 誤差を用いた価値関数の更新式は(式2.4)となる。 $\alpha$  は学習係数である。

$$V(s) = V(s) + \alpha(r(s, a) + \gamma V(s') - V(s)) \quad (2.4)$$

このような強化学習は固定的な環境での学習方法であるが、変動する環境下での強化学習も提案されている。

#### 2.3.2 強化学習の脳内機構

行動神経科学において、報酬は物体や事象に対する行動の頻度を増加させるものと考えられていた[55]。

このような脳内の報酬としてドーパミンが考えられている。Schultz はサルの報酬に対するドーパミンニューロンの発火を調べた[?]。ランプが点灯したときにレバーを押すとジュースがもらえることをサルに学習させた。学習前には、ジュースが与え

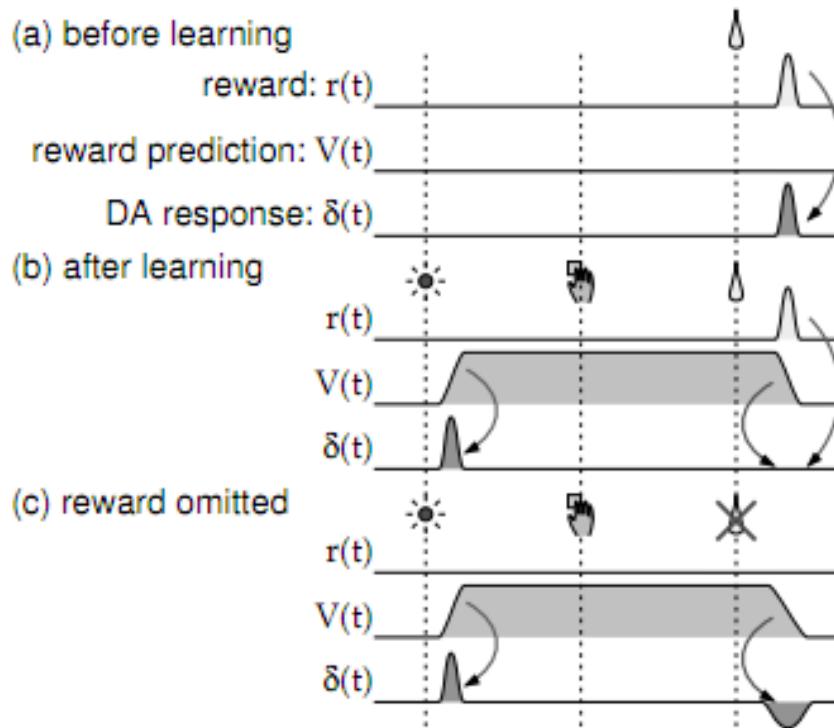


Fig.2.7 ドーパミン細胞の報酬予測応答 [67] .

られたときにドーパミンニューロンが発火する。しかし、学習後にはランプが点灯したときにドーパミンニューロンが発火し、ジュースが与えられたときには発火しなかった。ジュースが与えられなくなると、ジュースが与えられるはずのタイミングでドーパミンニューロンが抑制された。

これらのドーパミンニューロンの活動を TD 誤差としてとらえることができる [52, 17]。さらに、線条体ニューロンの活動も運動の結果得られる報酬予測と関係していると考えられている [18]。

ドーパミンニューロンと線条体の関係は Wickens らによって調べられている [38]。線条体のニューロンの Hebb 型の可塑性には、大脳皮質からの入力、線条体ニューロンの発火、ドーパミンの入力の 3 条件がそろったときにシナプスが増強される。ドーパミンが欠乏した状態で、大脳皮質からの入力に対して線条体ニューロンが発火すると、シナプスは弱化する。

これらの研究から、線条体で価値関数や行動価値関数が学習され、大脳皮質－基底核ループのダイナミクスで行動が選択されるという仮説が提案されている [4]。

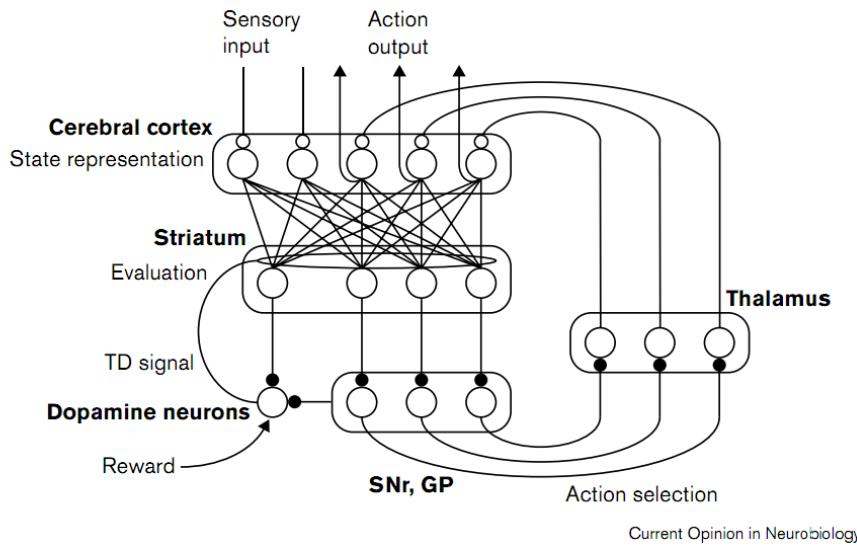


Fig.2.8 大脳基底核の TD 学習モデル [4] .

### 2.3.3 競合学習

強化学習は環境が変化しないことを前提としているが、実環境には様々な環境が存在する。そこで、学習するモジュールを複数用意し、環境を分節化しつつ学習する枠組みとして、競合学習が考えられている。この枠組みによって、環境が変化してもモジュールの切り替えや組み合わせにより、環境に適応することができる。

複数のモジュールによる制御は”mixture of experts”により導入された[19]。この機構は異なる専門をもつ複数のエキスパートと、用いるエキスパートを選択するゲートから構成されている。ゲートはエキスパートの出力を用いた分類器となっている。

次に、制御対象の予測器とコントローラの組みをもつモジュールが提案された[12]。予測誤差最小の予測器のモジュールを選択し、そのコントローラにしたがって制御、学習が行われる。しかし、最小のモジュールのみを選択する場合、学習初期における不安定な状態において、環境の分節化がうまく行われない。

そこで、予測誤差が最小のモジュールが最大の値をとる soft-max 関数により、予測値の合成、予測器の学習を行う概念的なモデルが提案された[60]。soft-max 関数は環境の分節化が進んでいない状況ではモジュール間の値の差は少ないが、分節化が進むにつれモジュール間の差が大きくなっていく。この soft-max 関数の値にもとづいて、システム全体の出力のための各モジュールの出力の重み付け、各モジュールの学習の重みが決定される。この機構は”Multiple Paired Forward-Inverse Models (MPFIM)”または”Modular Selection and Identification for Control (MOSAIC)”と呼ばれている。

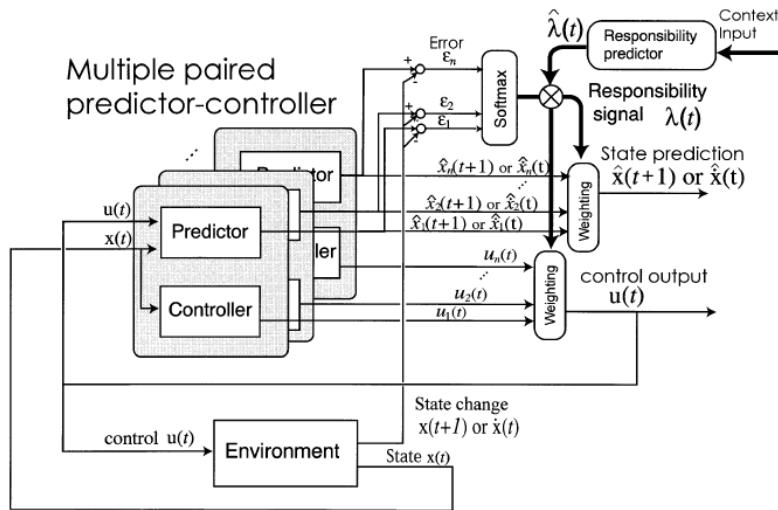


Fig.2.9 Multiple Paired Predictor-Controller architecture [13].

MOSAIC モデルを発展させたモデルとして、隠れマルコフモデルと EM アルゴリズムを用いたものや、強化学習と組み合わせた”Multiple Model-based Reinforcement Learning (MMRL)”などが提案されている [14, 13]。これらのモデルはコントローラ、学習手法以外は同一であるので、”Multiple Paired Predictor-Controller architecture”と呼ばれる。

### 2.3.4 脳のモジュール性

人が多数の道具の使い方を習得し、匠につかいながら生活している。新しい道具の使い方を習得するときより、すでに以前使ったことある道具を使うときの方が使いこなすのに時間がかかることが知られている。このことから、人間はすばやく道具が使えるように適応しているのではなく、人間の脳にはそれぞれの道具に対応するモジュールがあり、そのモジュールを切り替えることで適応的に行動できていると考えられている。

人間が新しい道具が使えるようになる過程での小脳の活動を観察すると、最初は広い範囲で見られていたのが道具が使えるようになると狭い範囲のみでしか活動が見られなくなった [11]。最初は広い範囲で学習が行われているが、学習の進行にともない小脳の特定の部分に内部モデルがつくられていることが示唆される。

複数の道具に対し使用しているときの脳の活動を記録すると、道具ごとに活動する場所が異なっていることが確認される [?]。道具の内部モデルが小脳にモジュールとして作られていることが分かる。

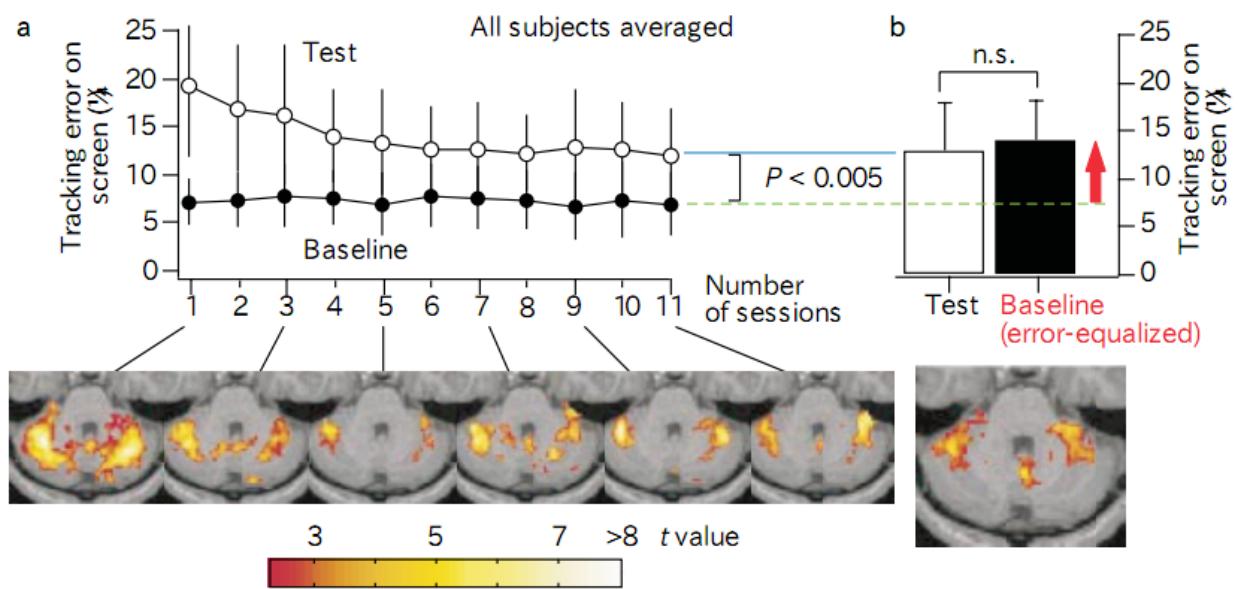


Fig.2.10 道具使用の習得と小脳の活動 [11].

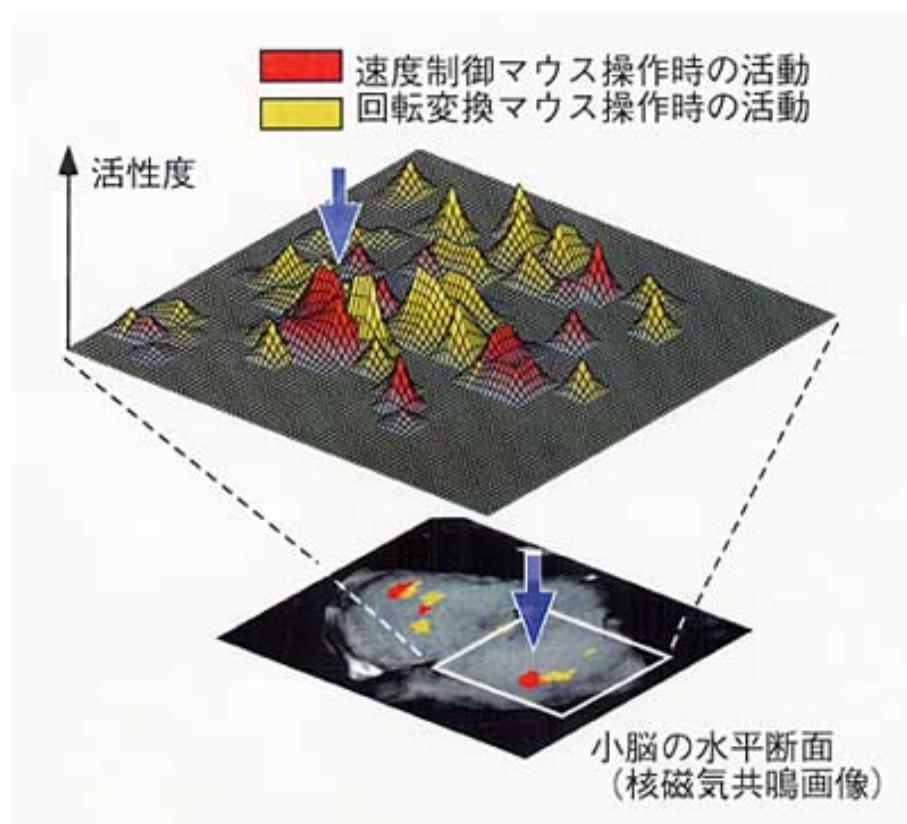


Fig.2.11 異なる道具使用時の小脳の活動 [63] .

### 2.3.5 exploitation and exploration dilemma

強化学習のような報酬を最大化する問題において、現在得られている知識を使って報酬が最大になるような行動を選択する。この行動選択の方法を *exploitation* とよぶ。また、得られている知識が不完全な場合や環境が変化する場合は、選択された行動が最適とは限らない。そこで、最適な行動を見つけるために、選択されない行動を選択することも重要である。このような探索を *exploration* とよぶ。

青斑核のノルアドレナリンニューロンが *exploitation* と *exploration* のバランスに影響を与えているのではないかと考えられている [15]。

## 2.4 乳児の探索行動と内発的動機づけ

乳児を特徴づける行動として、積極的に環境に働きかけ、探索的に周囲の環境を認識していく行動がある。この探索的な行動を引き起こす動機づけとして、内発的動機づけが考えられている。

### 2.4.1 心理学における内発的動機づけ

心理学において行動そのもの以外に報酬がない場合に内発的動機づけによる行動と呼ばれる。この定義では、子どもの遊びや探索的な行動は内発的動機づけによる行動と特徴づけられる [46]。

内発的動機づけを説明しようという最初の試みは、1950年代に人間の探索的行動を欲求が満たされていないのを軽減するためという動因低減説によるものだった [30]。しかし、刺激や情報を積極的に求めていく行動は決して欲求を満たすものではなく、動因低減説では説明できないとして批判された [59]。White はコンピテンスという概念を提唱した。コンピテンスとは自分の活動によって環境を思い通りに操作できる能力のことを指す。環境との相互作用の結果得られる効力感による動機づけはインフェクタンスと呼ばれる。

他の理論としては認知的不協和理論がある。これは、自己の内部にある矛盾する認知がある時に、不快感を感じ、その不快感を解消しようと生物は動機付けられるという理論である [25]。しかし、この理論も多くの人々の行動は認知の不一致を減少させるだけでなく、増加させようとする行動もあると批判された。Hunt は子どもを情報処理システムとみなし、最適水準の不適合を維持しようとして動機づけられると考えた [31]。このような考え方を発展させ、Berlyne は情報処理過程には不確かさを経験したとき、不確かさを軽減させようとする特殊的探索と、必要な刺激作用を獲得するため

の拡散的探索の2つ機制があると指摘した[2]。

Banduraは自己効力感という概念で学習理論を発展させた[1]。自己効力感とは自分が行為の主体として適切に行動することができるという確信である。自己効力感はある行動がどのような結果をもたらすかという結果予期、結果をもたらすための行動をどの程度うまくできるかという効力予期を区別して考える。

Vigotskyは現在の能力の範囲と、他者のわずかな支援によって達成できる範囲の間の領域を「発達の最近接領域」と定義した[56]。

#### 2.4.2 内発的モチベーションの脳内機構

脳内の報酬予測誤差を中脳のドーパミン細胞が表現している[23]。

ドーパミン系の機能は多様で複雑であることが分かっている。ドーパミン系の発火は新規な刺激や予期しない報酬を含む多くの事象が原因である[29]。一方で覚醒水準の低下や予測される興奮の減少によって抑制される[49]。ドーパミン系を人工的に刺激されると、周囲の環境を探索したくなり好奇心を示すようになる[35]。

強化学習の計算論ではドーパミンの発火をTD-errorとしてとらえている[52]。しかし、ドーパミン系は予期していないかった報酬に結びつかない刺激に対しても発火する[26]。そこで、ドーパミンの発火を驚きと重要性の組み合わせとしてモデル化するもある[8]。他のモデル化としては、ドーパミンニューロンの発火を予期していないかった重要な刺激へ注意を向けるものとして解釈したモデルもある[34]。

#### 2.4.3 内発的動機づけモデル

好奇心などの内発的動機づけは未知環境を探索し、環境のモデルを獲得することを目指す。このため、環境のモデルを獲得することを目的とした学習における探索の方針として好奇心のモデル化が考えられてきた。

コントローラーのアウトプット  $o_A(t)$  によって環境の状態が  $x(t)$  から  $x(t+1)$  に変化することを考える。この環境のモデルを獲得するために、現在の状況  $x(t)$  とアウトプット  $o_A$  から、次の時刻の状況  $x(t+1)$  を予測することを目指す。この時、好奇心は予測器の学習を促進させようとする欲求として定義される。

このような学習の手法として強化学習が用いられる。内部に構築されているモデルと実際の環境に差異がある状況が強化される。つまり、予測がうまくできない状況を作り出すようにコントローラーが行動を選択する。このことにより、環境の学習が進み予測ができるようになり、その行動は選択されなくなる。そのコントローラーがその状況に対して飽きている状態となる。

このときの問題として、本質的に予測できない環境を学習しようとすることがある。また、簡単な環境から学習して、その後に難しい環境を学習うして欲しい場合、学習の難易度の指標が必要である。この2つの問題を解決するためには、ともに学習の効果を学習する必要がある。そして、学習の効果の高い行動を選択することが望まれる。

ひとつの方法として、モデルの信頼度の変化の累積を最大化するものがある。信頼度を  $C(t)$  とすると(式2.5)の最大化としてシステムの目標を表すことができる。 $\gamma$  は割引率である。

$$E\left\{\sum_{t \geq t_0} -\gamma^{t-t_0} \Delta C(t)\right\} \quad (2.5)$$

他の方法として、モデルの各ステップの学習によって生じる出力の変化を学習する  $H$  を考える[47]。 $H$  は(式2.6)の近似となる。

$$E\{\Delta o(t)|i_t(t)\} \quad (2.6)$$

システムの目的は(式2.7)の最大化である。

$$E\left\{\sum_{t \geq t_0} -\gamma^{t-t_0} |H(t)|\right\} \quad (2.7)$$

他にも、予測誤差の小さいもとと大きいものの最適なバランスを保とうとするものや、予測誤差の減少を最大化するモデルがある[48, 16]。

このような行動選択の学習としては強化学習が有効である。報酬を実験者によって外的に決められるのではなく、自分で決める学習の仕方を内発的モチベーションによる強化学習と呼ぶ[5]。

#### 2.4.4 飼化

飼化は繰り返される刺激に対して、反応が弱まることである。新生児期にもある刺激が繰り返し提示されると、見る頻度や注視する時間が減少する飼化の現象がみられる。新生児にある刺激を繰り返し提示し注視する頻度が減少し基準を下回った後、他の刺激を見せると注視する時間が増加する脱飼化が見られた[7]。

Table2.2 感覚運動期の各段階 ([66] を参考とした)

段階の名称	時期	特徴
反射の行使	出生～1ヶ月	生得的な反射を行うだけ 反射も同化と調節により上達する
獲得性適応 第一次循環反応	1ヶ月～3.4ヶ月	反射的なシェマを他の対象にも適用し新しいシェマを獲得 新しく獲得されたシェマは何度も繰り返される 活動は自己の身体に限られている
第二次循環反応	3.4ヶ月～7.8ヶ月	意図的に外界と関わるようになる 複数のシェマを同時に用い、感覚的な運動を繰り返す
第二次シェマの協応	7.8ヶ月～1歳	隠した物を探したり、障害物を排除しようとしたりする 目的と手段が分離する
第三次循環反応 能動的実験	1歳～1歳6ヶ月	物を投げたりする行動を条件を変え繰り返し行う 実験により新しい手段を発見する
心的結合	1歳6ヶ月～2歳	失敗の連続の後に突然成功する 表象が出現する

## 2.5 認知発達に対するアプローチ

### ピアジェ理論

乳児の認知発達研究はピアジェによって初めて包括的な理論が提案された。

心理学者のピアジェは健常児を観察し、知能の発達段階を大きく2つに分類した。

1歳6ヶ月または2歳までを感覚運動的知能、それ以降を表象的知能と分類し、さらに感覚運動的知能6段階に、表象的知能を4段階に分類した[62]。乳児はピアジェの感覚運動的知能の時期にあたる。ピアジェの感覚運動期の段階と行動、認知の特徴をTable2.2に示す。

ピアジェは乳児期を感覚運動に基づく認知機能が発達していくと考えた。運動を反復的に行い適応させる「調節」と、ある運動を他の場面にも利用する「同化」によって乳児の運動は発達していく、環境との関係を運動によって理解するようになる。

### ダイナミック・システムズ・アプローチ

人間の発達における自己組織化を2つの仮定を中心に説明する。一つ目は発達する組織は多くの構成要素からなる複雑な組織であり、複雑な環境の中で環境に開かれた状態であるという仮定。二つめは、振る舞いの変化は様々な時間幅で起きるという仮

定。これによって、行動の発達のみではなく様々なものの発達を論じようという考え方である [50]。

### 認知発達ロボティクス

人の認知プロセスの解明を認知科学、発達心理学、神経科学などのかかわりの深い多くの分野の知見を使い、ロボットを用いた仮説と検証の繰り返しによって実現を目指す構成論的アプローチをとる [24]。認知プロセスは身体と環境のインタラクションから生じるという考えは、これまでの認知科学での体を出力機器とみなす考えを変えた。

### 本研究のアプローチ

これらの知見から、乳児は内発的動機づけに基づき、まだ学習していないが学習できそなものを選択的に探索することで、学習を効率的に行っていると考えられる。そして、自己の身体を学習の起点とし、そこから環境へ探索の範囲を広げていくことで、自己を対象とした運動から環境を対象とした運動へと運動が遷移していくと考えられる。この過程は状況の切り分けを導いていると考えられる。

そこで、本研究では感覚運動予測の予測誤差に基づく内発的動機づけによる運動生成によって、自己を対象とした運動から環境を対象とした運動へと運動が遷移が観察されるか検証する。



## 第3章

# 感覚運動予測に基づく 運動生成モデル

本章では、感覚運動予測に基づく運動生成モデルを説明する。

### 3.1 システムの構成

前章で紹介した知見から、システムに必要な構成を考察する。

周囲の環境を学習しつつ、適切に運動選択ができるようになるシステムを提案する。学習として乳児は運動と感覚刺激との関係性を学習していくことから、学習機構を感覚運動予測としてモデル化する。感覚運動予測と運動生成を同時に学習していくシステムを設計する。

システムに必要な要素として3つ考えられる。

- 感覚運動予測と運動生成の競合学習
- 内発的動機づけに基づく運動の重みの更新
- モジュールを統合する運動選択

感覚運動予測と運動は状況に応じて区別され、適切に行なうことが求められる。状況ごとに応じて、予測器と運動生成の組みのモジュールを複数もつシステムとして競合的に学習していくモデルとする。

運動生成は内発的動機づけに基づいて行なう。予測誤差から運動の重みづけを更新するものとしてモデル化する。

運動選択では、複数のモジュールを統合し、状況を判断し運動選択を行い運動指令を出力する。

全体のシステム概要を Fig.3.1 に示す。

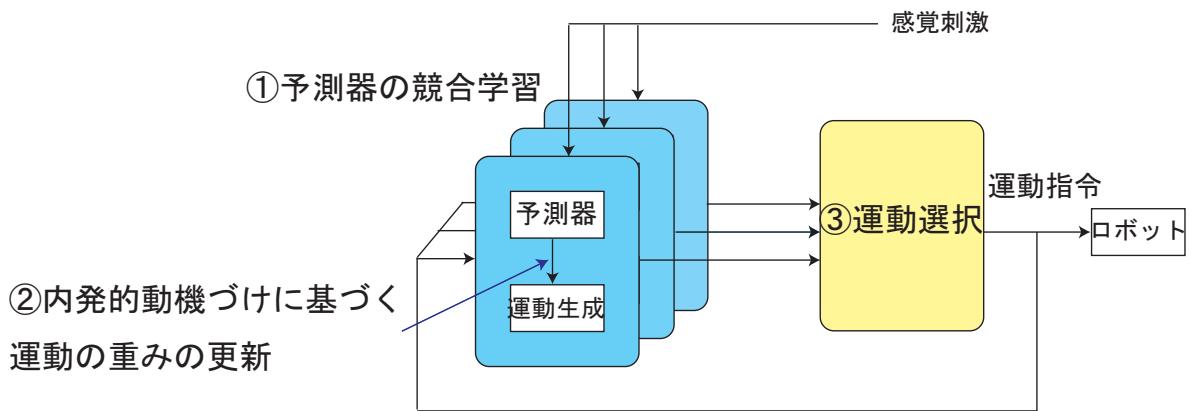


Fig.3.1 システム概要.

## 3.2 モジュール制御方式による競合学習

様々な状況に対応し適切な行動を選択する運動生成方式として，予測モデルと制御器の組をもつ複数のモジュールを用意し，予測誤差が一番小さいモジュールの運動制御器に従い運動生成を行うモジュール制御方式が提案されている [12] .

提案モデルでは，内発的動機づけに従って運動の重みを更新する．予測器と運動生成に関する運動の重みの組を持つモジュールを複数用意し，予測器を学習しつつ運動の重みを更新していく．

### 3.2.1 予測器の設計

運動によって環境に働きかけることが可能で，センサによって環境からの刺激を知覚できるエージェントを仮定する．このエージェントはある運動  $a$  を行ったとき，環境からの刺激  $r$  を知覚する．このとき，ある時刻  $t$  における運動  $a_t$  から次の時刻  $t+1$  における刺激  $r_{t+1}$  を予測する．予測関数を  $r = f(a)$  とし，運動と刺激の組合せからこの予測関数を学習していく．

### 3.2.2 運動の重みの更新

すべての運動に対し重みが 0 の状態から運動するごとに重み  $W(a)$  を更新する．

運動の重みは内発的動機に基づき，予測誤差によって加算していく．選択された運動以外も運動の類似度を考慮して興味度を増加させる．

増加量  $\delta$  はステップ状の関数として (式 3.1) によって定義する． $e$  は予測誤差， $\eta, \alpha, \beta$  は任意の定数である．

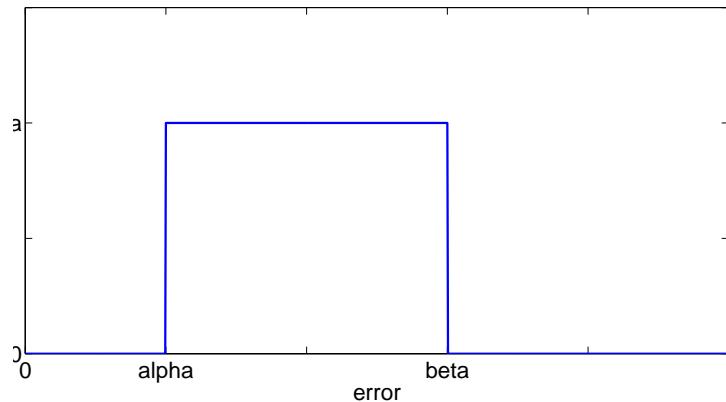


Fig.3.2 予測誤差に対する運動の重みの加算値.

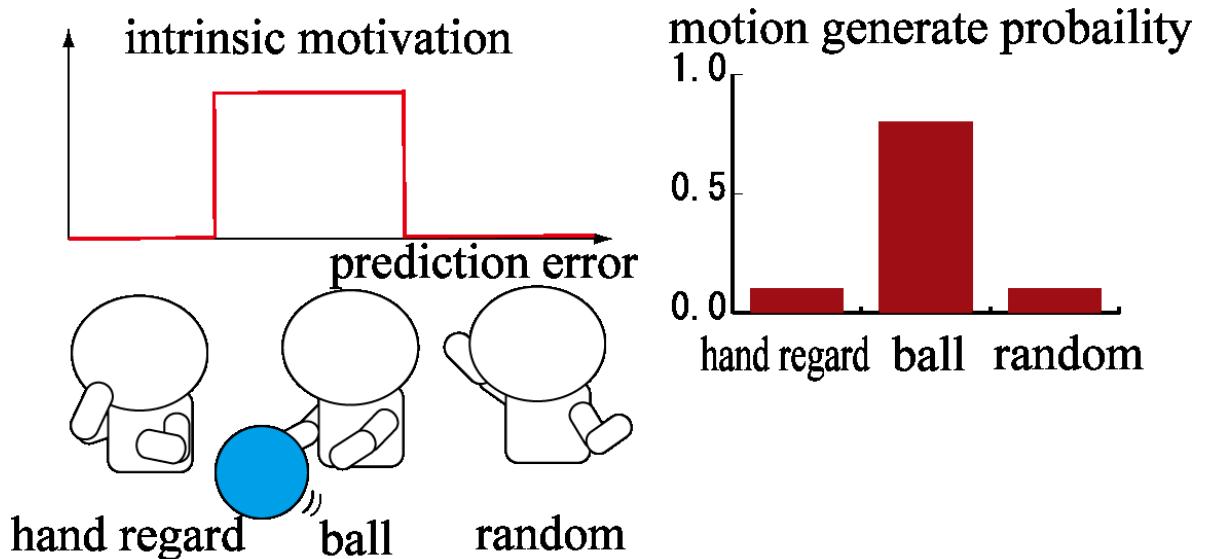


Fig.3.3 獲得される運動の重み.

$$\delta = \begin{cases} \eta & (\alpha < e < \beta) \\ 0 & (e \leq \alpha, \beta \leq e) \end{cases} \quad (3.1)$$

グラフに表すと 3.2 となる .

運動の予測誤差に対し , 時間が経過したあとに獲得される運動の重みを Fig.3.3 に示す .

#### 運動の類似性に基づく重みの更新

実際に行なった運動と類似の運動も運動の類似度に基づいて重みを増加させる . 運動の類似度は運動開始時と目的の手先の位置の近さで定義する .  $x_{pre}$  を運動開始時の

手先位置， $x_{dst}$  を目的の手先位置とする。手先位置による運動  $a$  の表現を  $x = \begin{pmatrix} x_{pre} \\ x_{dst} \end{pmatrix}$  とし，運動  $a_1$  と  $a_2$  の類似度  $R$  は(式3.2)で定義する。 $\gamma$  は任意の定数である。

$$R = \exp(-\gamma \|x_{a1} - x_{a2}\|^2) \quad (3.2)$$

の増加  $\Delta W$  は増加量  $\delta$  と類似度  $R$  の積として  $\Delta W = \delta \times R$  とする。

重みは常に増加するので，重みの総和が閾値  $W_{th}$  を越えたら和が閾値  $W_{th}$  となるように標準化する。

$$W_{t+1}(a) = \begin{cases} W_t(a) + \Delta W & (W_{sum} \leq W_{th}) \\ (W_t(a) + \Delta W) \times W_{th}/W_{sum} & (W_{sum} > W_{th}) \end{cases} \quad (3.3)$$

### 3.3 モジュールの統合と運動選択

ある時刻  $t$  におけるエージェントの状態を  $s(t)$  とする。時刻  $t$  での状態  $s(t)$  から時刻  $t+1$  での状態  $s(t+1)$  へ一定時間で遷移し，その後静止する。この状態遷移をエージェントの運動  $a$  とする。ある状態  $S_t$  から選択できる運動の集合を  $A_t = a_{t1}, a_{t2}, \dots, a_{tn}$  とする。 $n$  は状態  $S_t$  から選択できる運動の総数である。このとき，次の状態として同じ状態は除外する。

システム全体の運動選択の基準となる運動生成分布  $M(a)$  を各モジュールの運動の重みをもとに更新される。状態  $s_t$  における次の運動は，運動生成分布  $M(a)$  を最大とする運動が選択される。ただし，重みが最大の運動が複数ある場合はランダムに運動を選択する。

$$a = \arg \max_{a_i \in A_t} M(a_i) \quad (3.4)$$

運動生成分布は運動の重みと馴化によって更新される。用いられる重みは選択されたモジュールの重みである。

#### 順化

乳児は同じ行動には飽き，同じ行動をしにくくなる。馴化は微分方程式で表現される[51]。

運動生成分布  $M(t)$  は運動の重み  $W(t)$  を用いて差分方程式(3.5)で記述される。 $S(t)$  は馴化による項， $\zeta$  は任意の定数である。

$$\tau \Delta M(t) = \zeta(M(t) - W(t)) - S(t) \quad (3.5)$$

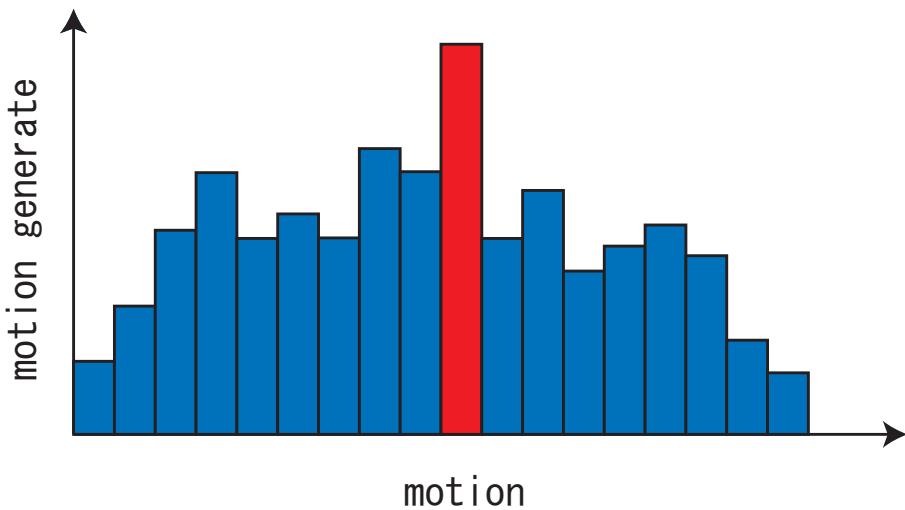


Fig.3.4 運動生成分布と運動生成.

$S(t)$  として , 時刻  $t$  に運動  $a$  を行なった場合は 1 , 行わなかった場合は 0 とする .

$$S_i(a) = \begin{cases} 1 & (if a(t) = a_i) \\ 0 & (otherwise) \end{cases} \quad (3.6)$$

### 馴化モデルの検証

(3.5) 式の馴化モデルの検証をおこなう . 3.5 には  $\tau = 10.0$  ,  $\alpha = 0.5$  ,  $1 \leq t \leq 19$  ,  $40 \leq t \leq 59$  ,  $80 \leq t \leq 99$  のとき  $S(a) = 1$  , それ以外のとき  $S(a) = 0$  として ,  $M(a)$  をプロットした .

3.6 には一度運動したときの馴化と脱馴化の様子をプロットした .  $t = 1$  , のときに  $S(a) = 1$  とそれ以外のとき  $S(a) = 0$  とした .

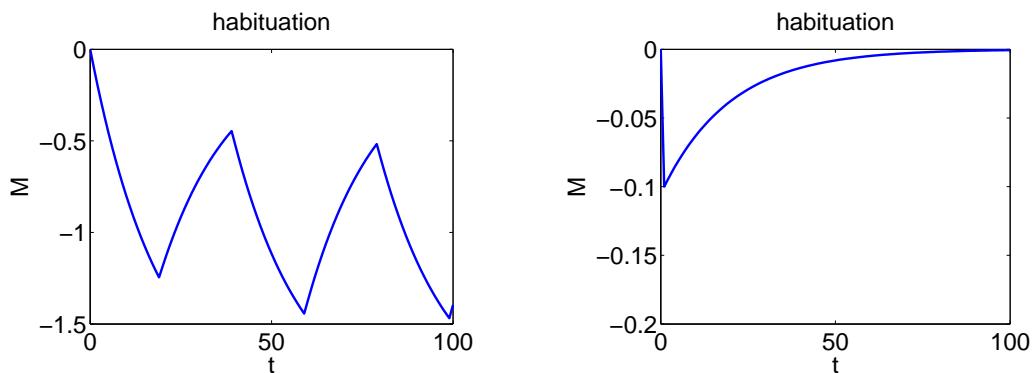


Fig.3.5 馴化による運動生成分布の変化

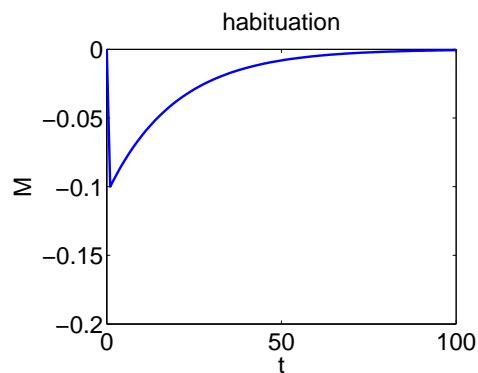


Fig.3.6 1回の馴化による影響

本章で述べたシステム構成を Fig.3.7 に示す .

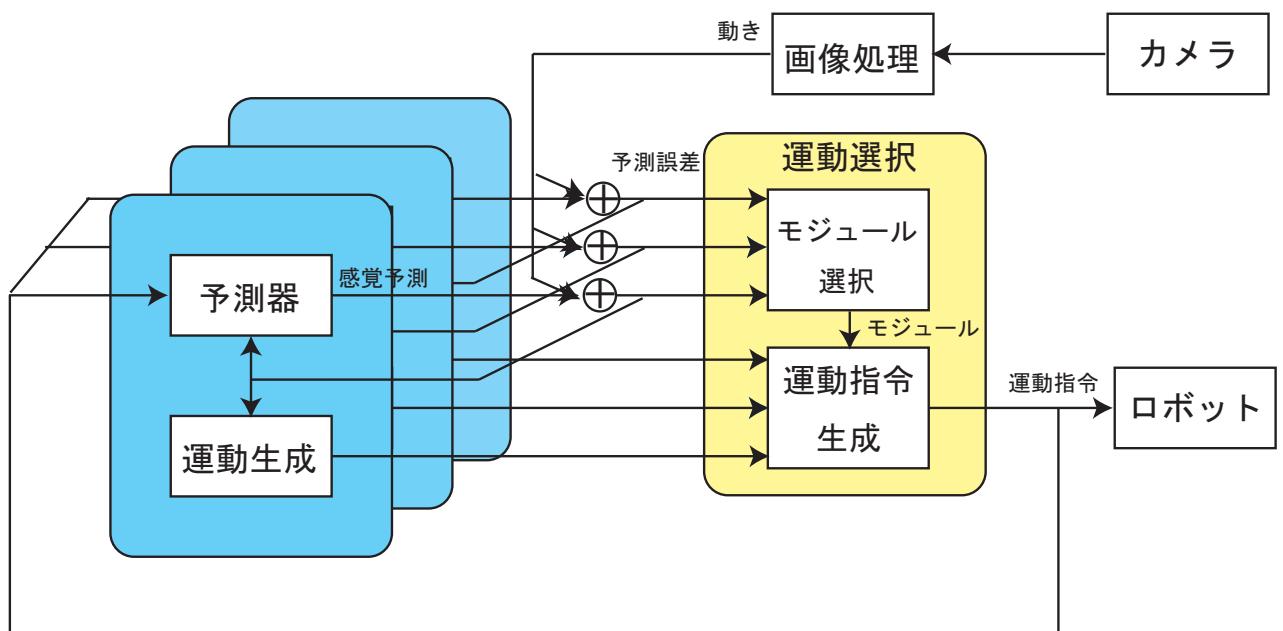


Fig.3.7 システム構成.

## 第4章

# 実験

本章では、前章で示したシステムの検証実験と実験の結果について述べる。

### 4.1 システム構成

前章で示した運動遷移についての実験を行うためのロボットシステムを構築する。プラットホームとして Aldebaran Robotics 社製ヒューマノイドロボット”NAO”を用いる。

#### ロボットシステム

高さ 573mm、幅 311mm、奥行き 275mm、重量 5.2kg の全身 25 自由度からなるヒューマノイドロボットである [39]。腕は 6 自由度からなり、それぞれの関節の配置を Fig.4.1、可動角を Table4.1 に示す。制御には Aldebaran Robotics 社による API の”NAOqi”を利用する [42]。”NAO”的制御方法は各関節の目標角度を絶対角度と到達時間を指定する。

#### ビジョン

画像は NAO の額にあるカメラを利用する。カメラから  $640 \times 480$  の解像度の画像を 1 秒間に 30 フレーム取得することができる。視野は  $72.6^\circ$  DFOV( $60.9^\circ$  HFOV, $47.6^\circ$  VF) である [41]。

Table4.1 Joint Range of NAO.

Joint Name	Right Arm		Left Arm	
	Min Angle(rad)	Max Angle(rad)	Min Angle(rad)	Max Angle(rad)
ShoulderPitch	-2.09	2.9	-2.09	2.9
ShluderRoll	-1.65	-0.09	0.09	1.65
ElbowYaw	-2.09	2.09	-2.09	2.09
ElbowRoll	0.09	1.56	-1.56	-0.09
WristYaw	-1.82	1.82	-1.82	1.82
Hand	0.00	1.00	0.00	1.00

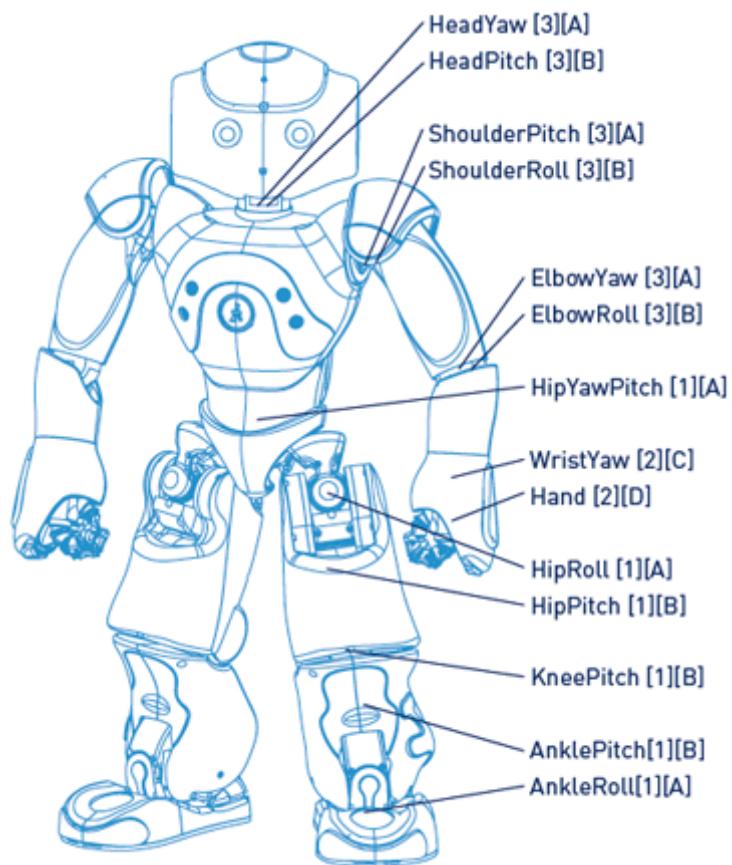


Fig.4.1 layout joint of NAO[39]

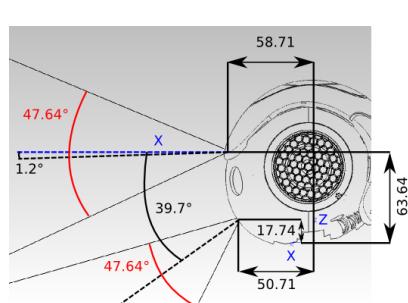


Fig.4.2 Camera position lateral

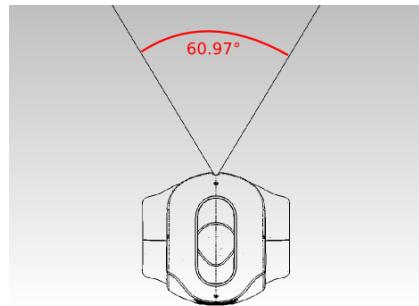


Fig.4.3 Camera position top

Table4.2 Used joint and joint angle.

Joint name	Joint angle		
LShoulderPitch	0.0	0.3	0.6
LShluderRoll	-0.2	0.1	0.4
LElbowRoll	0.0	0.6	1.2

## 4.2 実装

### 4.2.1 モジュールの競合学習

モジュールは運動するごとに予測誤差最小のものに切り替え，学習していく．モジュールの数を本研究では 20 とする．

### 4.2.2 運動生成

ロボットの状態を関節角の組み合わせで表現し，その状態の遷移を運動とする．運動は "AOqi" の数関数 "angleInterpolat" によって絶対角と遷移にかかる時間を指定して生成する．遷移にかかる時間は 1 秒とした．

NAO の 3 つの関節 "LShoulderPitch", "LShoulderRoll", "LElbowRoll" を使いそれぞれ角を 3 つの値に量子化した．つまり，27 の状態があり，運動は全部で 702 となる．それぞれの関節で使用した関節の値を Table4.2 に示す．

#### 運動生成のパラメータ

(式 3.1) で表される運動の重みの増加量  $\Delta K$  のパラメータとして， $\eta = 0.01$ ， $\alpha = 20$ ， $\beta = 60$  とする．

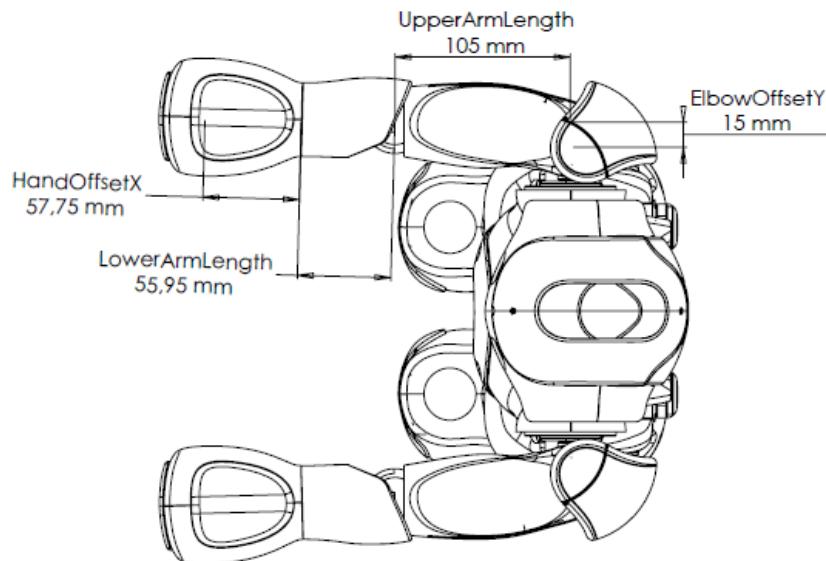


Fig.4.4 Length of links

運動の重みの総和の閾値は  $W_{th} = 1$  とする .

(式 3.5) の運動生成分布の馴化のパラメータとして ,  $\zeta = 0.5$  ,  $\tau = 10.0$  , とする.

$$\Delta M(t) = 0.05 \times (M(t) - W(t) - 1) \quad (4.1)$$

運動生成分布と運動の重みの更新は , 学習初期の不安定な状態を避けるため , 運動を 50 回経過してから行う .

#### 運動の類似度

運動の類似度を計算するために上腕の長さ 105mm , 前腕の長さ 55.95mm , 手の長さ 57.75mm を用いた [40] . (式 3.2) のパラメータ  $\gamma$  は 1000 とした .

### 4.2.3 感覚運動予測に用いる手法と特徴量

#### 感覚運動予測の手法

感覚運動予測の手法として線形回帰を用いる . 回帰行列を  $A$  とし , 入力ベクトルを  $x$  , 出力ベクトルを  $y$  とするこれらとの関係は (4.6) となる . 入力ベクトル , 出力ベクトルはそれぞれ縦ベクトルである .

$$y = Ax \quad (4.2)$$

この回帰行列をのために , 最近 100 組の入出力の組み合わせから  $A$  を求める . 入力ベクトルを転置し , 行方向に並べたものを  $X$  , 出力ベクトルを転置し , 行方向に並

べたものを  $Y$  とする。これらの関係は (4.7) と表せる。

$$Y = XA^T \quad (4.3)$$

これらから  $A$  は (4.4) のように求めることができる。

$$A = (X^{-1} Y)^T \quad (4.4)$$

逆行列の計算には OpenCV の逆行列計算 `cv::Mat::inv` を用いた。特異値分解により擬似逆行列を求める関数である。

#### 予測器の初期値

予測器の回帰行列の初期値として、行列の各成分にランダムな値を与え初期値とした。

#### 4.2.4 感覚運動予測に用いる特徴量

感覚運動予測では運動指令から運動運動によって起こるオプティカルフローを予測する。

##### 運動指令

運動は関節角を指定することで行われる。運動指令としてある時刻での状態の関節角と、次の状態の目標の関節角を並べたベクトルとする。関節は 3 つ使うので、6 次元のベクトルとなる。

##### オプティカルフロー

画像からオプティカルフローを求めるために OpenCV の関数 `"cv::calcOpticalFlowFarneback"` を用いた。Gunnar Farneback のアルゴリズムを用いてオプティカルフローを計算する関数である。画像を  $3 \times 3$  に分割してここで求められたオプティカルフローの大きさの平均をとる。 $3 \times 3$  のオプティカルフローの大きさを並べることで 9 次元のベクトルとし、オプティカルフローのベクトルとする。

オプティカルフローは fps が大きい方が安定して求めることができる。そこで、運動開始から 0.3 秒おきにオプティカルフローを計算し、一つの運動で 1.5 秒後まで 5 回計算する。この 6 回のオプティカルフローの総和を予測する特徴量として用いる。

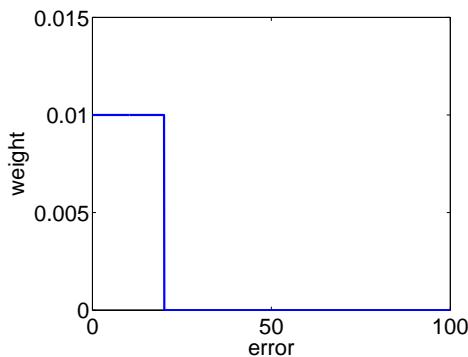


Fig.4.5 小さい予測誤差に対する動機づけ

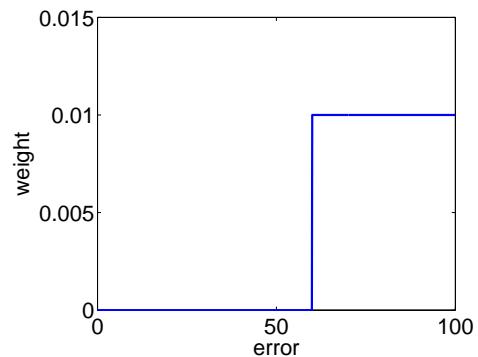


Fig.4.6 大きい予測誤差に対する動機づけ

### 4.3 実験

以上のシステムを用いて実験を行った。

実環境の状況は様々に変化し、周囲に物体がある場合もあればない場合もあると考えられる。しかし、この様な変化する状況では運動の遷移を定量的に扱うことが困難である。そこで、常に物体が手の届くところにある状況を作り運動の遷移についての実験をおこなう。常に手が届く範囲にボールがある状況での運動遷移について実験をおこなった。

固定した状況では異なる運動により得られる感覚刺激の複雑さや刺激が質的異なることはあるが、特定の運動に対する応答は変化せず状況の切り分けはおこなわれないと考えられる。周囲に物体が状況とない状況をつくり、獲得される状況の切り分けについて検証する。ボールに手が届く範囲にある状況とない状況を交互に作り、獲得される状況の切り分けを考察した。

それぞれの実験では内発的モチベーションに基づく運動生成の他に、異なる動機づけに基づく運動生成を行い比較した。異なる動機づけとしては、小さい予測誤差の運動に対して運動の重みが増加するモデル、大きい予測誤差に対して運動の重みが増加するモデルを用いて比較した。それぞれの動機づけの予測誤差に対するグラフを Fig.4.21, Fig.4.26 に示す。

### 4.4 運動遷移実験

ボールを NAO の手の届くところに置いた状態で運動生成をおこなうことで運動遷移についての実験を行った。



Fig.4.7 ボールを用いた実験の様子

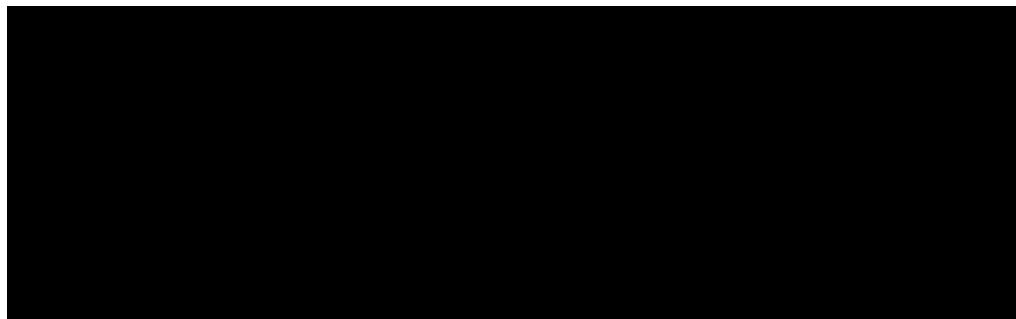


Fig.4.8 運動の生起頻度

NAO の前にサッカーボールを置いた状態で NAO を動かし，ボールを動かす運動，ボールを動かさず視野中に手を動かす運動について運動の頻度を求め，運動の遷移を検証した．

#### 4.4.1 運動遷移の評価

運動の遷移はある時間幅での中心的な運動のカテゴリの変化で評価する．中心的な運動のカテゴリはすべての運動を等確率で生成した場合比べ，そのカテゴリの運動の生起頻度が十分に高いカテゴリとする．運動の生起頻度は運動 100 回における各カテゴリの運動をした数とする．

Table4.3 各カテゴリの運動数.

カテゴリ名	運動数	全運動に対する割合
全運動	702	1.0
ボールを動かす運動	162	0.23
視野中に手を動かす運動	296	0.42

### 等確率での運動の生起頻度

それぞれのカテゴリに当てはまる運動数と全運動に対する割合を Table4.3 に示す。ボールを動かす運動は実際に全運動をさせたときにボールが動くか観察し、視野中に手が写る姿勢は NAO のカメラからの映像を見て判断した。

### 運動遷移の評価方法

等確率に運動生成をおこなった場合における各カテゴリの運動の生起頻度について標準偏差を計算した。標準偏差は視野中に手を動かす運動は 3.7353、ボールを動かす運動は 5.0972 である。

正規分布を仮定した場合、平均  $\pm$  標準偏差の値となる確率は 68.26% であり、平均 + 標準偏差より大きくなる確率は 15.87% である。あるカテゴリの運動の生起頻度が平均 + 標準偏差より大きくなつたとき、運動はそのカテゴリの運動に偏っていると考え、そのカテゴリの運動が中心的な状態であると考えられる。

また、観測値が平均に対してどのくらい大きいかは Z 値で評価できる。 $Zscore > 1$  となる運動を中心的なカテゴリとする。この中心的なカテゴリの遷移で運動遷移を評価する。

カテゴリ  $i$  の生起頻度  $Freq_i$  に対する Z 値  $Z_i$  は等確率で運動した時の平均  $mean_i$  と標準偏差  $std_i$  を用いて(式 4.5)で計算できる。

$$Z_i = (Freq_i - mean_i) / std_i \quad (4.5)$$

### 4.4.2 内発的動機づけモデル

内発的動機づけモデルにより運動生成をおこなった際の運動の生起頻度を調べた。

ボールを動かす運動、手を視野中に動かす運動について運動 100 回ごとにそれぞれ何回おこなわれるか調べ 4.12 に示す。運動遷移分布を更新する前のランダムに 50 回運動させた後を 0 とした。

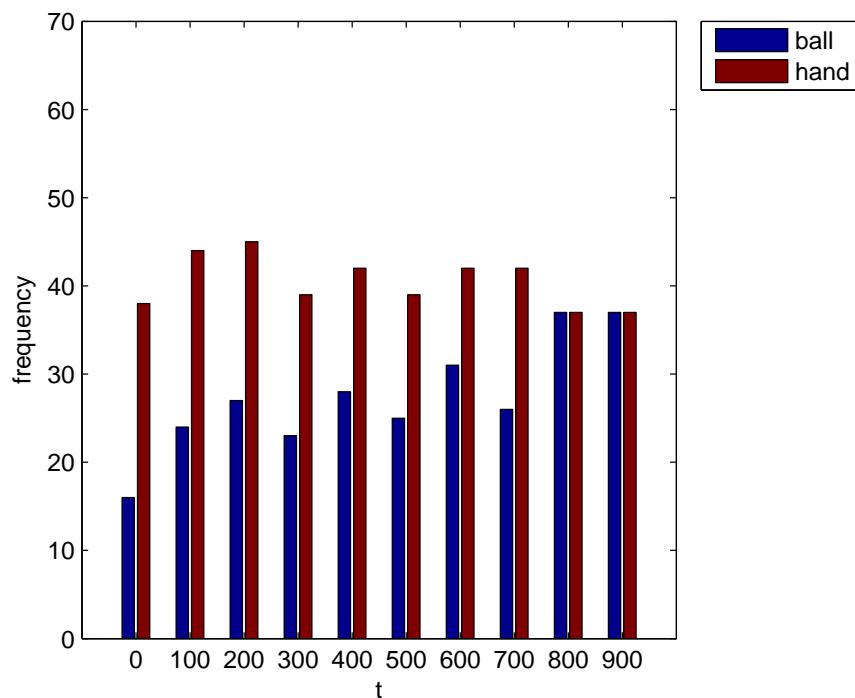


Fig.4.9 内発的動機づけモデルによる運動の生起頻度

視野中に手が写る運動は 200-300 回でもっとも高い値をとり，それまでは頻度が増加し，その後減少する傾向が見られる．視野中に手を動かす運動の全運動に対する割合は 0.42 であり， $100 < t \leq 300$  ではその値を越えている．

ボールを動かす運動は計測終了まで次第に増加していく傾向が見られる．ボールを動かす運動はの全運動に対する割合は 0.23 であり， $t > 500$  ではその値を越えさらに増加していることが分かる．

運動の偏りは予測誤差に基づく運動の重みの更新により，選択されやすい運動ができることで生じる．そこで，時間経過とともにそれぞれの運動の予測誤差の変化と，重みを増加させる範囲となった運動の回数を調べた．

予測誤差の時間変化を見ると手が写る運動に比べ，ボールを動かす運動は誤差が減少するのに時間がかかることが分かる．時間経過とともに重みを増加させる運動の中で，視野中に手を動かす運動の割合が高い状態から，ボールを動かす運動の割合が高い状態へ遷移することで，全体の運動が視野中に手を動かす運動に偏っている状態から，ボールを動かす運動が偏る状態へ遷移すると考えられる．

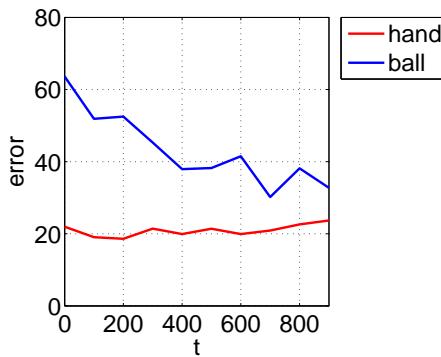


Fig.4.10 内発的動機づけモデルによる予測誤差の時間変化 .

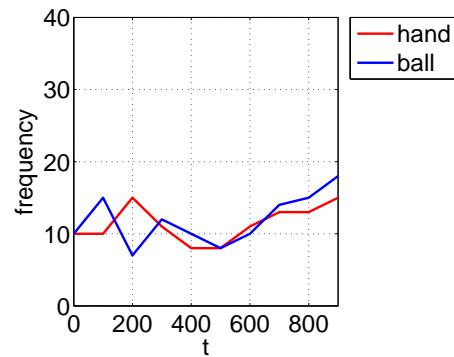


Fig.4.11 内発的動機づけモデルによる運動の重みが増加した回数の時間変化 .

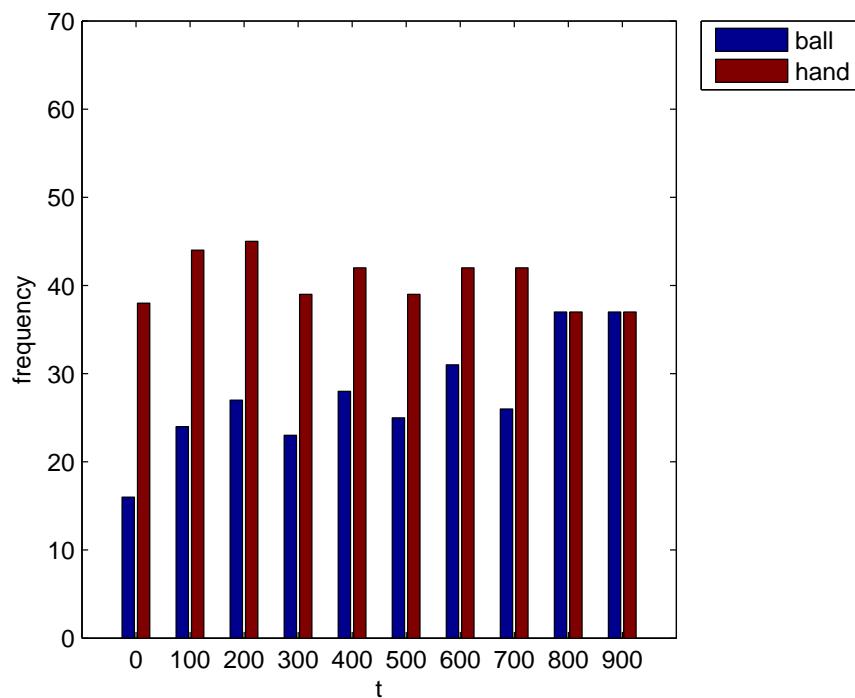


Fig.4.12 内発的動機づけモデルによる各カテゴリの Z 値 .

### 運動遷移の評価

実験で得られた各カテゴリの運動の生起頻度から Z 値を求めグラフにした .

$0 < t < 300$  では視野中に手を動かす運動の Z 値の方が高く , それ以降はボールを動かす運動の方が Z 値が高くなっている . 視野中に手を動かす運動からボールを

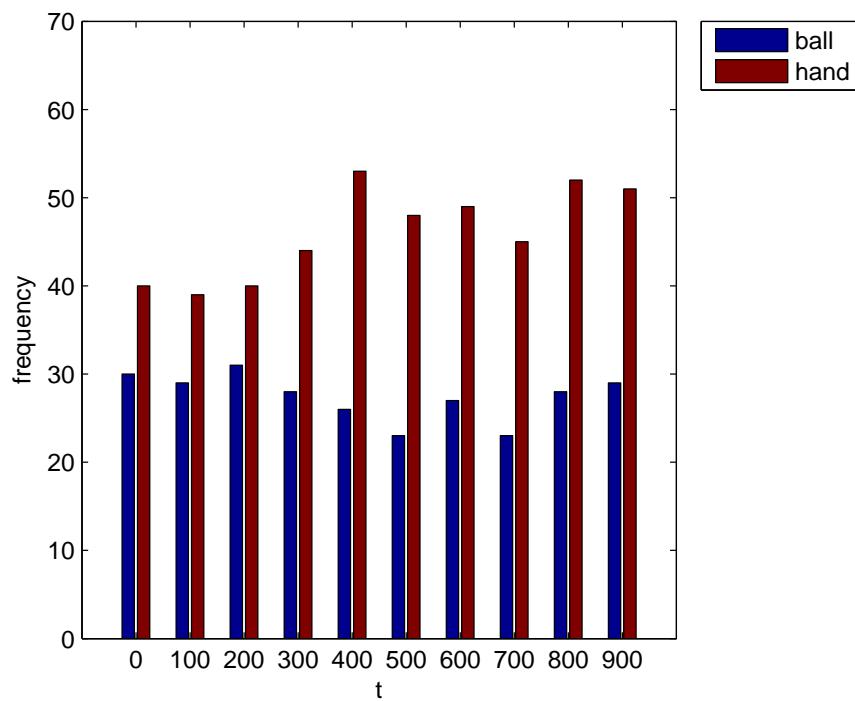


Fig.4.13 学習初期状態での運動の生起頻度

動かす運動へ運動が遷移する傾向が見られる。しかし、視野中に手を動かす運動において $Z$ 値は1以下であり有意に生起頻度が高くなっているとはいえない。

本研究のシステムは、運動の重みの増加により運動が偏るシステムとなっているが、それと同時に予測器も学習し重みを増加させる運動は変化していく。今回の実験で視野中に手を動かす運動に偏る状態となっていても、予測器の学習が速く生起頻度として現れない可能性がある。

あるの時刻での中心的なカテゴリがどのカテゴリか評価するために、予測器を学習初期の状態と、学習後期の状態に固定した上でどのような運動の偏りが生じるか検証した。

#### 学習初期での運動の偏り

提案モデルの実験での100回運動後に獲得された予測器を用い、予測器は学習せずに予測誤差に基づく興味度の増加だけをおこなった。その時の各カテゴリの運動の生起頻度をFig.4.13に示す。

ボールを動かす運動は20回後半あまり変化していない。視野中に手を動かす運動は300回経過した頃から増加し、50回程度の生起頻度となっている。

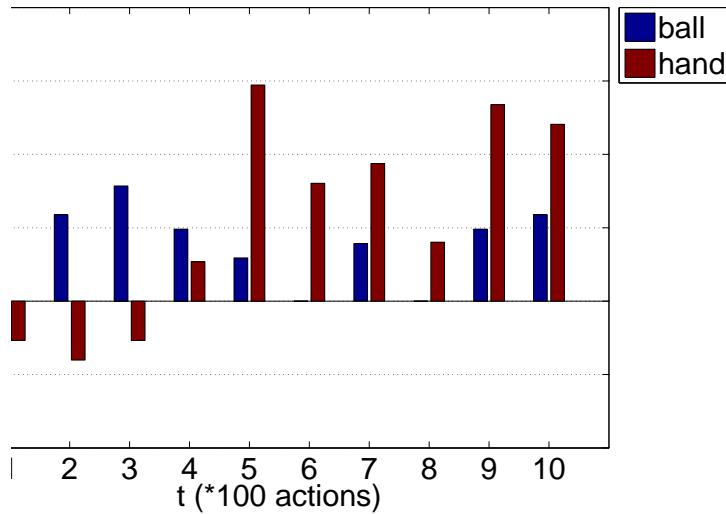


Fig.4.14 学習初期状態での Z 値

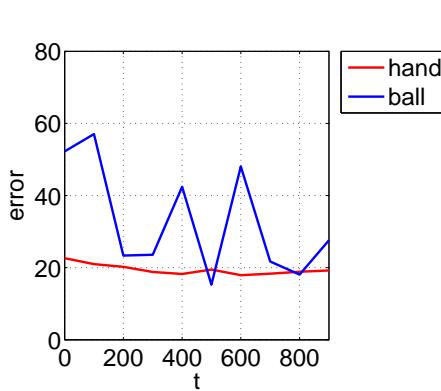


Fig.4.15 予測誤差の時間変化

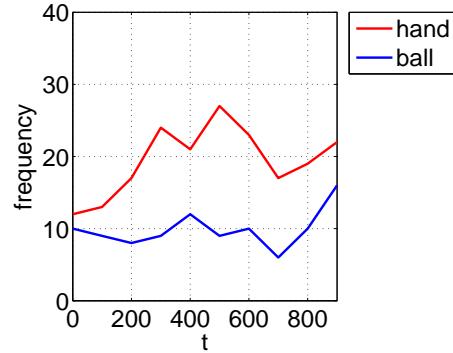


Fig.4.16 興味度が増加した運動の回数の時間変化

運動頻度の Z 値を Fig.4.14 に示す。

$t \leq 3$  ではボールを動かす運動のみが Z 値が 1 をこえている。 $5 \leq t \leq 7, 8 \leq t \leq 10$  では視野中に手を動かす運動のみが Z 値が 1 をこえている。

運動生成開始したばかりのときは、運動選択が等確率に近く、全体的傾向とは異なる運動の重みが偶然増加することが起こる。運動生成後半に視野中に手を動かす運動の Z 値が 1 より大きくなっていることから、学習初期の段階ではボールを動かす運動をよく行う状態であるといえる。

視野中に手を動かす運動の予測誤差は約 30 であり、ボールを動かす運動は約 60 となっている。興味度が増加する範囲の運動は実験中常に視野中に手を動かす運動の方がボールを動かす運動よりも多く、視野中に手を動かす運動が選択されやすくなっている。

いることが分かる。

運動の偏りの  $t$  検定 それぞれのカテゴリを等確率で動かした場合と比べ、有意に運動が増加しているか  $t$  検定を用いて評価した。

各カテゴリの生起頻度の平均が、等確率に動かしたときの生起頻度の平均より有意に増加しているかを、各カテゴリの生起頻度から等確率に動かしたときの生起頻度を引いたものの平均が 0 であるという帰無仮説のもと、平均が 0 より大きいという対立仮説に対して有意水準 5% で  $t$  検定をおこなった。

視野中に手を動かす運動では  $p$  値が 0.2838 となり帰無仮説を棄却できなかった。

次に等確率に運動したときに、視野中に手を動かす運動の理論的な生起頻度の平均である 42 回より有意に平均が大きいという対立仮説に対して、有意水準 5% で  $T$  検定をおこなった。 $p$  値は 0.0179 となり帰無仮説は棄却され有意に平均が 42 回より大きいことがいえる。

学習初期の段階では視野中に手を動かす運動の頻度が増加することが示された。

ボールを動かす運動についても同様に  $t$  検定をおこなった。ランダムに動かした場合より学習初期の状態の運動の生起頻度の平均が増加しているという対立仮説に対して有意水準 5% で  $t$  検定をおこなった。

$p$  値は 0.0264 となり、帰無仮説は棄却された。学習初期の段階においてボールを動かす運動の生起頻度は増加しているといえる。

$t$  検定ではどちらのカテゴリもランダムより生起頻度が大きくなっているが、その原因として標本数が少ないと、運動生成開始後に偶然ボールを動かす運動に偏ったことが原因と考えられる。

#### 学習後期での運動の偏り

学習後期として、提案モデルの 800 回運動したときの予測器を用いて、予測器を学習させずに運動生成をおこなった。そのときのそれぞれの運動の生起頻度を Fig.4.17 に示す。

視野中に手を動かす運動は生起頻度が増加していき、300 回以降は 30 回を少し越えるくらいとなっている。

$Z$  値を Fig.4.18 に示す。

$4 \leq t \leq 9$  においてボールを動かす運動のみ  $Z$  値が 1 を超える状態となっている。運動生成が始まった直後は手を動かす運動に偏っている傾向が見られるが、運動生成が安定する後半ではボールを動かす運動の  $Z$  値が高くなっていることから、学習後期ではボールを動かす運動に偏る状態であるといえる。

手を視野中に移動させる運動の予測誤差は 20 を下まわり、ボールを動かす運動は

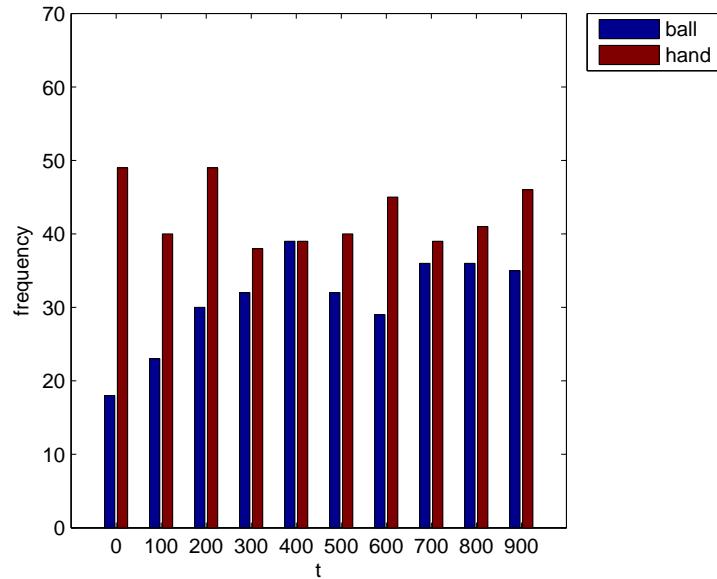


Fig.4.17 学習後期での各カテゴリの運動の生起頻度

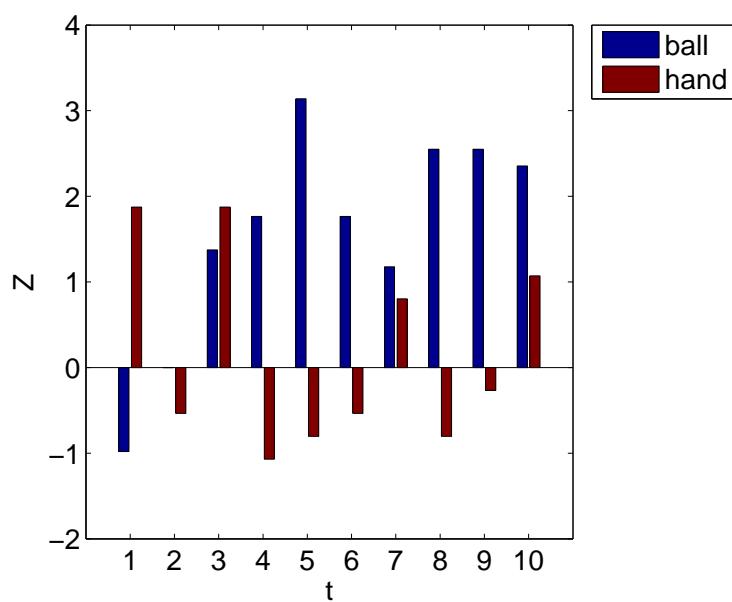


Fig.4.18 学習後期での各カテゴリのZ値

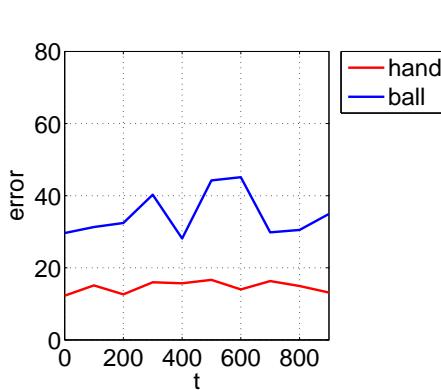


Fig.4.19 予測誤差の時間変化

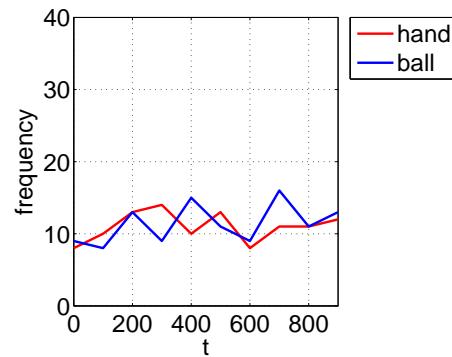


Fig.4.20 興味度を増加させる運動の回数の時間変化

30から40の範囲となっている。興味度を増加させる運動は実験中常にボールを動かす運動の方が、視野中に手を動かす運動よりも多かった。学習後期ではボールを動かす運動選択されやすくなっていることが読み取れる。

運動の偏りのt検定 それぞれの運動が等確率に動かしたときと比べて、有意に増加しているかt検定をおこなった。

等確率で動かしたときと視野中に手を動かす運動の平均は等しいという帰無仮説の下で、ランダムに動かしたときの平均は学習後期の運動生成の平均より大きいという対立仮説に対して有意水準5%でt検定をおこなった。p値は0.9534となり、帰無仮説は棄却されなかった。視野中に手を動かす運動の生起頻度が増加しているとはいえない。

等確率に運動させたときとボールを動かす運動の平均は等しいという帰無仮説の下で、ランダムに動かしたときの平均は学習後期の運動生成の平均より大きいという対立仮説に対して有意水準5%でt検定をおこなった。p値は $4.3037 \times 10^{-4}$ となり、帰無仮説は棄却された。ボールを動かす運動の生起頻度は有意に増加しているといえる。

学習後期の運動生成ではボールを動かす運動のみ生起頻度が増加したといえる。

#### 4.4.3 異なる動機づけモデルとの比較

興味度の増加させる範囲によって運動遷移が異なり、乳児の発達にみられる運動遷移が提案モデルによる興味度の増加のさせ方によって現れることを検討する。予測誤差が小さい運動の興味度を増加させるモデルと、予測誤差が大きい運動の興味度を増加させるモデルによる運動生成をおこなった。興味度の増加量を定義する提案モデル

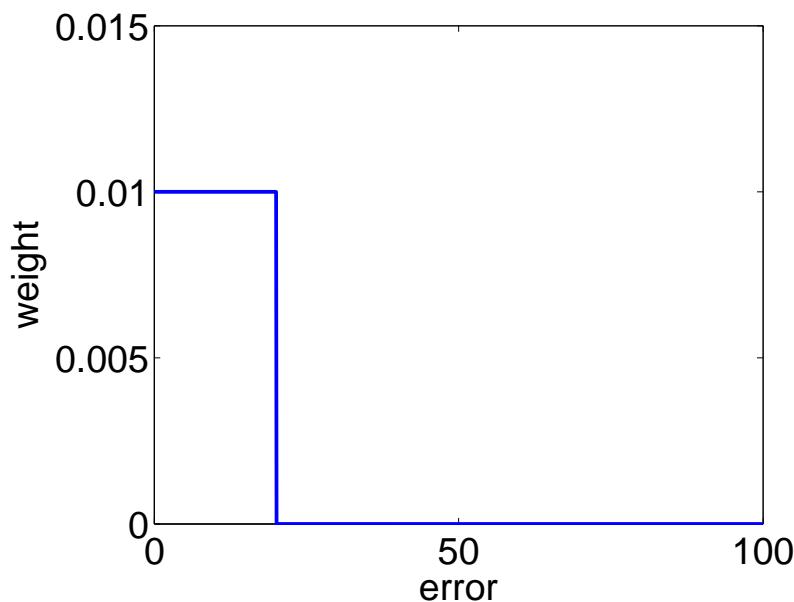


Fig.4.21 予測誤差の小さい運動に対する重みの増加モデル

の(式3.1)を変更したモデルを用いて実験をおこなう。

興味度の増加範囲以外の影響をできるだけ小さくするため、予測器の初期値と運動生成分布を更新する前の50回ランダムに運動させるところまでは、提案モデルの実験のデータを用い、その後から実験を開始した。

#### 予測誤差の小さい運動に対する運動の重みの増加モデル

予測誤差の小さい運動の重みを増加させるモデルの重みの増加量は(式4.6)で定義する。

$$\delta = \begin{cases} \eta & (e \leq \alpha) \\ 0 & (e > \alpha) \end{cases} \quad (4.6)$$

このモデルを用いて運動生成実験をおこなった。運動の生起頻度をFig.??に示す。視野中に手を動かす運動は徐々に増加していき約60回で安定となっている。ボールを動かす運動は約15回で変化はない。

興味度を増加させる運動の回数は視野中に手を動かす運動の割合が大きいことが分かる。

運動遷移の評価 各カテゴリの運動の生起頻度におけるZ値をFig.4.25に示す。

実験中常に視野中に手を動かす運動のZ値のみが1より大きくなっていることが分かる。常に視野中に手を動かす運動をしやすい状態であり、運動の遷移が起こって

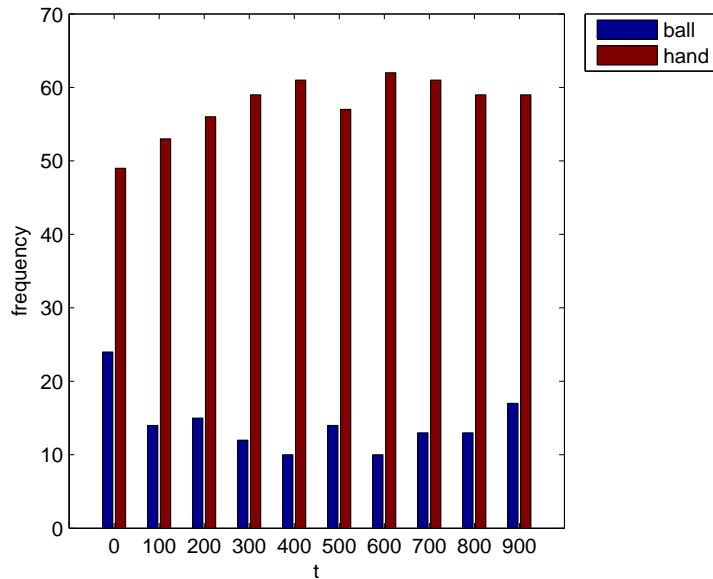


Fig.4.22 運動の生起頻度

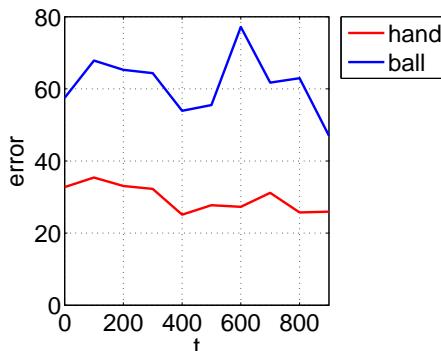


Fig.4.23 予測誤差の時間変化

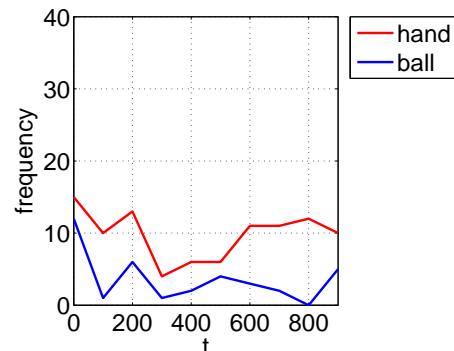


Fig.4.24 興味度が増加した回数の時間変化

いないうことが分かる。

#### 予測誤差の大きい運動に対する興味度の増加モデル

予測誤差の大きい運動の興味度を増加させるモデルの興味度の増加量を(式4.7)で定義する。

$$\delta = \begin{cases} \eta & (e \geq \beta) \\ 0 & (e < \beta) \end{cases} \quad (4.7)$$

予測誤差の大きい運動に対して興味度を増加させるモデルを用いて運動生成をおこ

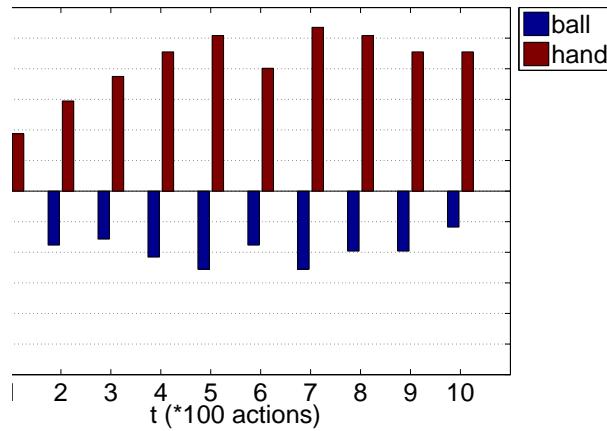


Fig.4.25 小さい予測誤差に対する重みの増加モデルの Z 値

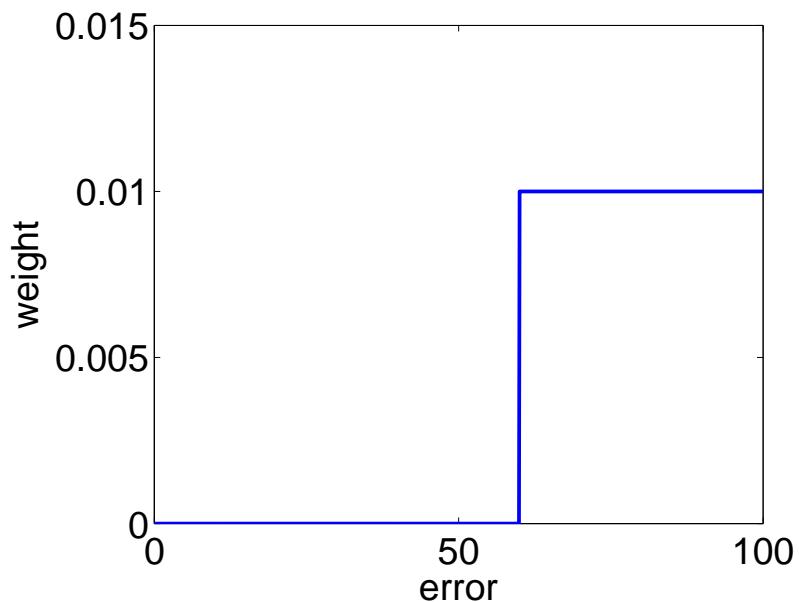


Fig.4.26 予測誤差の大きい運動の重み增加モデル

なった。それぞれのカテゴリの運動の生起頻度は 4.30 に示す。

視野中に手を動かす運動、ボールを動かす運動はともに約 35 回で変化はない。

重みを増加させる運動はボールを動かす運動の割合が大きいことが分かる。

運動遷移の評価 各カテゴリの運動の生起頻度における Z 値を Fig.?? に示す。

常にボールを動かす運動の Z 値のみが 1 より大きいことが分かる。常にボールを動かす運動をしやすい状態であり、運動が遷移することはなかった。

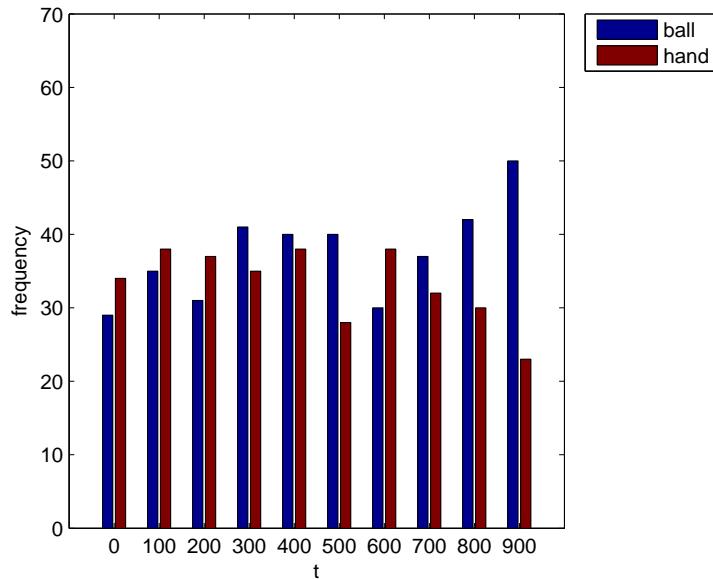


Fig.4.27 運動の生起頻度

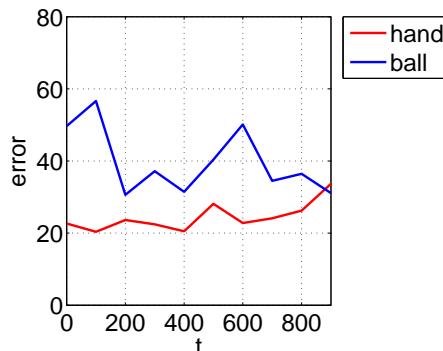


Fig.4.28 予測誤差の時間変化

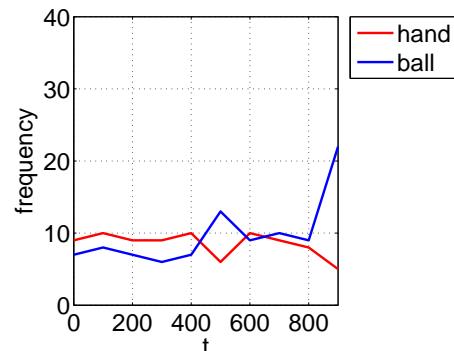


Fig.4.29 重みが増加した回数の時間変化

#### 4.4.4 運動遷移実験に関する考察

学習初期の段階において視野中に手を動かす運動とボールを動かす運動の両方で運動の生起頻度が大きくなったのは、今回の予測器では視野中に手を動かす運動の学習はより早い段階で収束し、視野中に手を動かす運動とボールを動かす運動の両方の興味度が増加したことが原因だと考えられる。

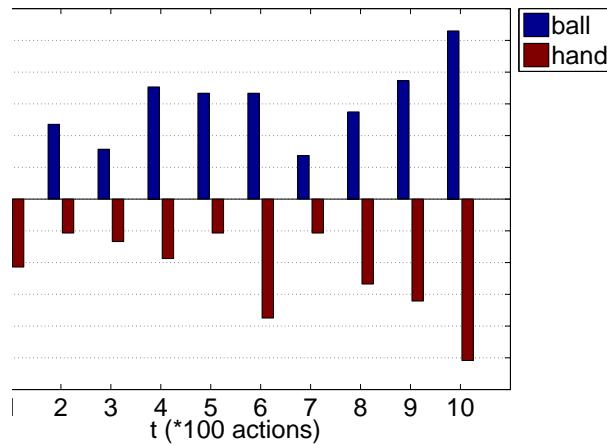


Fig.4.30 予測誤差が大きい運動に対する重みの増加モデルにおける Z 値

Table4.4 運動の生起頻度と予測誤差の相関係数と p 値.

ボールを動かす運動の予測誤差		視野中に手を動かす運動の予測誤差	
相関係数	p 値	相関係数	p 値
ボールを動かす運動	0.1308	0.5155	0.5101
視野中に手を動かす運動	-0.5165	0.0058	-0.3259

### 運動の偏りと予測誤差

運動遷移実験をおこなっている中で、運動が偏らない運動は予測誤差が小さくならず学習が進まないことが観察された。そこで、運動の生起頻度と予測誤差の平均の相関係数と、相関なしという仮説を検証するための p 値を計算した。

ボールを動かす運動の生起頻度と視野中に手を動かす運動の予測誤差、視野中に手を動かす運動の生起頻度と視野中に手を動かす運動の予測誤差は、直感を支持する結果となった。ボールを動かす運動の生起頻度とボールを動かす運動の予測誤差には p 値が大きく有意な相関関係は認められなかった。視野中に手を動かす運動の生起頻度とボールを動かす運動の予測誤差は直感に反する結果となった。

視野中に手を動かす運動はボールを動かす運動の約倍の運動の数がある。視野中に手を動かす運動へ運動が偏った場合、極端にボールを動かす運動の回数が減り、直感に反する結果となった可能性がある。

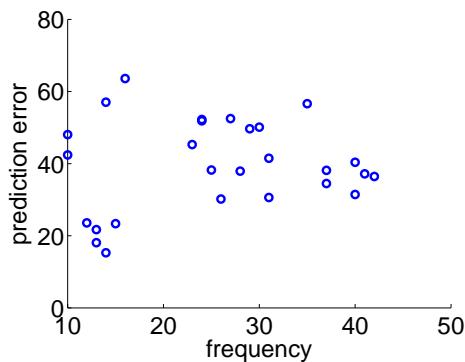




Fig.4.35 ボールがない状態の実験の様子



Fig.4.36 ボールがある状態の実験の様子

#### 4.5.1 状況の切り分けの評価方法

外部の状況が内部で区別されているかを検証する。状況の切り分けができるることは、状況ごとに予測器が区別されていることに対応する。ボールを動かす運動のカテゴリにおいて、ボールがある状況とない状況で選択される予測器の回数を調べた。状況の切り分けを評価する指標として以下の2項目が考えられる。

1. 状況によって選択される予測器の異なる度合い
2. 予測の正確性

状況によって選択される予測器の異なる度合いとして、以下の2項目が挙げられる。

1. 各状況において使用される予測器の乖離度
2. 1つの状況に使用される予測器のばらつきの少さ

**乖離度の指標** 状況によって予測が分かれているとき、各予測器を使用する回数の差が大きくなると考えられる。そこで、予測器の使用された回数の差の2乗和を乖離度として用いる。乖離度を(式4.8)に示す。 $div$ は乖離度、 $Freq_{i,j}$ はカテゴリ*i*におけるモジュール*j*の選択された回数、 $n$ は各カテゴリの実験に使用した運動の回数である。

$$div = \left\{ \sum_j (Freq_{1,j} - Freq_{2,j})^2 \right\} / 2n^2 \quad (4.8)$$

**ばらつきの指標** ばらつきの指標として平均情報量を用いる。平均情報量は(式4.9)で定義される。

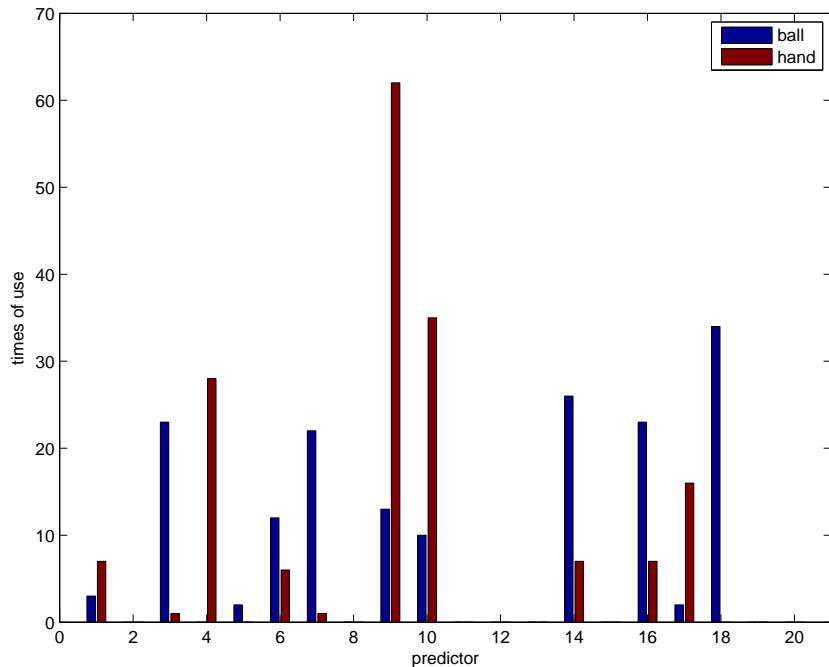


Fig.4.37 提案モデル

$$H(P) = - \sum_{A \in \Omega} P(A) \log(P(A)) \quad (4.9)$$

平均情報量が大きいほどばらつきが大きいことを示している。予測器  $j$  が用いられる確率を  $P(j) = Freq_j/n$  として平均情報量を求める。平均情報量をばらつきの指標として用い、ばらつきの少なさを平均情報量の低さで評価する。

#### 4.5.2 状況の切り分けの評価

それぞれの獲得された予測器に対し、ボールを動かす運動とボールがない状態での同じ運動で予測をおこない、用いられる予測器を調べ状況の切り分けができているか評価した。

各モデルにおける乖離度を Fig.4.40 に、ばらつき度合いを Fig.4.41 に、予測誤差を Fig.4.42 に示す。

乖離度、ばらつきの小ささの両方の指標において、大きい予測誤差により重みが増加するモデルが状況の切り分けをよくできていることを示している。予測の正確性では提案モデルの予測誤差が小さく、もっとも予測できていると考えられる。

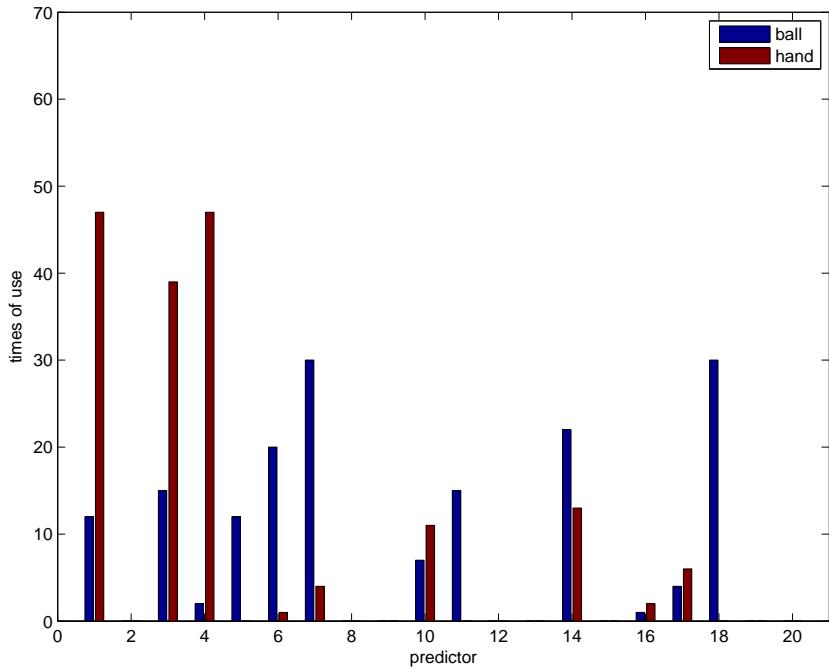


Fig.4.38 小さい予測誤差により興味度が増加するモデル

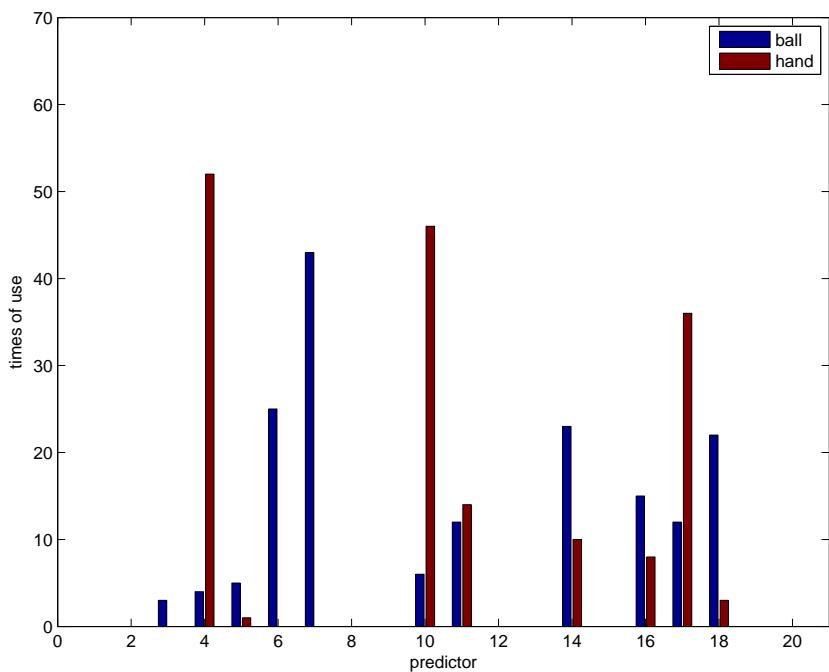


Fig.4.39 大きい予測誤差により興味度が増加するモデル

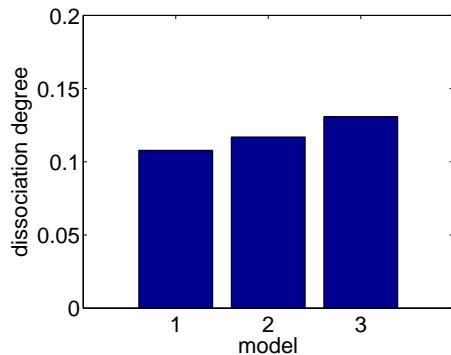


Fig.4.40 各モデルの乖離度

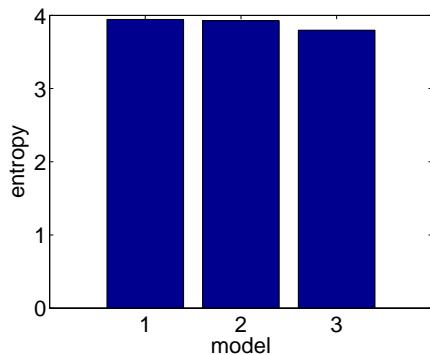


Fig.4.41 各モデルのばらつき

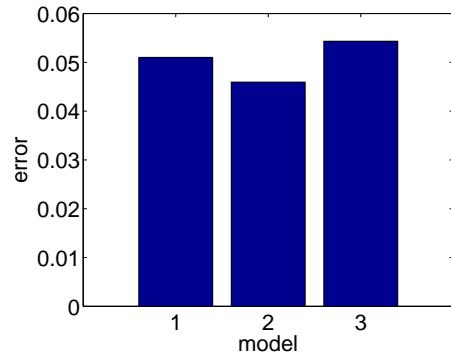


Fig.4.42 各モデルの予測誤差

- 1: 特殊的探索
- 2: 特殊的探索 +  
拡散的探索
- 3: 拡散的探索

### 4.5.3 状況の切り分けについての考察

状況の切り分けは大きい予測誤差に対して興味度を増加させるモデルがもっとも状況の切り分けができていた。学習できていない運動に対して積極的に学習するようになることが、状況の切り分けを促進していると考えられる。

予測の正確性は内発的動機づけによるモデルが最も予測誤差が小さく、すでに学習が進行しているものをさらに学習すことと、新しいものを学習しにいくことのバランスが学習の進行に重要であることが分かる。

状況の切り分けを評価する2項目のバランスから、内発的動機づけモデルが状況の切り分けを最も促進するといえる。

また、今回の実験設定では将来的に学習可能なものしかない状態であった。学習不可能なものがある環境では、大きい予測誤差の運動ばかりおこなうと、学習が進まない可能性がある。このような状況では内発的動機づけモデルの最適さがより顕著とな

ると考えられる。

## 第5章

### 結論と展望

本章では、結論と展望を述べる。

#### 5.1 結論

乳児の運動発達において自分の手をよく見る自己を対象とした運動から、ボールで遊ぶような環境を対象とした運動への運動遷移が観察されている。しかし、この運動発達の原理は解明されていない。さらに、この運動発達の順序は様々な環境を適切に区別しながら学習するときに重要な状況の切り分けを促進させると考えられる。

本研究では内発的動機づけに基づく運動生成によって、乳児において観察される自己を対象とした運動から環境を対象とした運動への運動遷移を説明することが出きることを示した。自己の身体から環境への運動の遷移は、単純なものから優先的に学習していくというメカニズムを外部から観測したときに見られる運動遷移であることが確かめられた。

また、内発的動機づけにより自己の能力を徐々に拡大していくように探索していくことは、適切な状況の切り分けを導き、乳児の認知発達をガイドする役割を担っている可能性が示唆された。

運動の偏りは、特定の運動の生起頻度をあげ、学習の妨げとなる可能性のある他の運動の影響を受けにくくし、学習を加速させていると考えられる。しかし、運動が偏りすぎると他の学習可能なものの学習を妨げる結果ともなりかねない。内発的動機づけは運動の偏りによる学習の促進と、他のものへ学習の範囲を広げていくことの適切なバランスを保っており、乳児の認知発達を適切に導いていると考えられる。

## 5.2 展望と課題

提案モデルでは、自己の身体を対象とした運動から環境中の物体を対象とした運動への遷移を実現した。乳児は自己と物体、あるいは自己と他者との1対1の関係から、物体を介した他者とのインタラクションへと発達していく。このような他者とのインタラクションも自己から物体、物体から他者へと統一的に発達していくと考えられる。このシステムは物体を対象とした運動から、さらに複雑な運動へ遷移していく、他者とのインタラクション、他者との物体を介したやりとり遊びなどに発展していくことができる可能性がある。

本研究では感覚運動の関係性の違いから、状況の切り分けが起きるようになることを示した。状況は環境との関係性で記述することができる。区別された状況を環境の要因に帰属することで、環境の分類ができるようになり表象的知能に結びつくと考えられる。

また、乳児は発達とともに自己と環境を区別することができるようになり、環境にも物体と他者が存在することに気づくようになる。自他分離の最初の段階として、自己と環境の区別があげられる。乳児は環境を知覚すると同時に自己についての認知を深めていく。本研究の状況の切り分けはこの自己と環境の区別を確立するまでの最初の段階となると考える。

区別された状況に基づいてその要因を環境中の物体に帰属していく、自己と環境を区別できる認知システムを構築することができると考える。

### 5.2.1 自己形成の段階

自己形成の段階における主な考えをまとめると、

#### ナイサーの自己概念

ナイサーは自己の概念を5つに分類した[33]。

- 生態学的自己 (ecological self): 知覚される身体的自己
- 対人的自己 (interpersonal self): 他者との社会的交渉に基づく自己
- 概念的自己 (conceptual self): 自分自身についての言語化された心的表象
- 持続的自己 (temporally extended self): ライフ・ストーリーとしての自己
- 神秘的自己 (private self): 自分は他人とは違うことへの気づき

### Rochat の自他分離の発達段階

Rochat は自他分離の発達段階を 6 段階に分類した [43] .

level 0 : 混同 鏡に映った像に無関心で , 鏡自身にも無関心である状態 . 鏡に映った像と鏡が映している実際の環境を混同する .

level 1 : 区別 鏡に映った像と実在するものとの区別ができる段階 . 運動の感覚と完全に随伴する視覚刺激があることが理解できる .

level 2 : 状態 鏡に映った運動と固有感覚の感覚間の関係性を探索することができる . 混同ではなく , 鏡に映った像は自分だと気づくことができる .

level 3 : 同一化 鏡に映った像は自分だと認識することができる . 頭にシールを貼った状態で鏡の前に座らせると鏡に手を伸ばすのではなく , 自分の頭に手を伸ばしてシールを取ろうとする . 自己の概念が出現する段階である .

level 4 : 永続性 自己の永続性が現れる段階 . 映像が明らかに若かったり , 場所が違ったり , 着ている服が違ったりしていても , その映像が自分だと認識できる .

level 5 : 自己意識 , メタ自己意識 第三者からどう自分が見えるかを認識できる . 他人による評価を気にするようになる .

上の分類で混同の段階にあたる新生児でも自分の体が他の存在と違うことが分かっていることを示す結果がある . 生後 24 時間の新生児の頬に指が触れたときの反応は , 実験者が触れたときの方が新生児自身で降れたときよりも強く反応した [44] .

こうした自己と環境を区別する能力は自己の身体だけに特別な知覚経験によると考えられる . 自分の泣き声であったり , 自分の体を触ったときの感覚であったり , 自分の腕が視野を横切ったりしたときには , 手が顔に触れていると同時に顔が手に触れている感覚を知覚する . 自分の手を顔に持っていく行動は生まれた直後や妊娠期間の最後の 3 ヶ月によく見られる .



## 謝辞

本研究は、東京大学工学部機械情報工学科國吉・原田研究室において、國吉教授のご指導のもとで行われました。本研究を支えてくださった研究室外の方々に、こころより感謝致します。

指導教員の國吉康夫教授には、研究室配属以来、親身で熱心なご指導をいただきました。研究室配属時から認知発達研究における考え方や、発表の仕方、論文の書き方に至るまで丁寧にご指導をいただきました。ありがとうございました。

原田達也准教授からは、研究に関する深いご理解とご支援をいただきました。また、研究会などで有益な御指摘を頂いた大村吉幸助教に感謝致します。研究室の技官、秘書の方々には、事務手続きから研究のサポートまで、学生生活を送る上でさまざまな面倒を見て頂きました。深く感謝いたします。

研究の先輩方には、研究を進める上で大変多くの手助けをいただきました。特に、修士課程の山田康智先輩には、研究テーマ決め、モデルの設計、論文の書き方など、さまざまな相談をさせていただきました。尾形邦裕先輩、崎朝子先輩には、研究会や発表練習で親身に自分の発表を聞いていただき、熱心な指導をしていただきました。鷺坂隆志先輩には、部屋のミーティングの際に研究方針に関して熱心にアドバイスをいただきました。栗山貴嗣先輩、佐藤宏樹先輩には研究テーマ決めの際に有益なご意見をいただきました。塚原祐樹先輩には、自分でも整理しきれていない考えを簡潔にまとめていただき、発表の仕方についてもアドバイスをいただきました。丹野雅之先輩には、テーマ決めや発表練習の際に研究の問題点に関して指摘していただきました。中村皓佑先輩、西川銳先輩はいつも気を使って声をかけてくださいました。山下裕也先輩には、グラフの書き方、引用の仕方を教えていただき、引用の修正もしていただきました。佐々木諒介先輩、志田和也先輩には積極的に研究の相談にのっていました。図や動画の作成を手伝っていただきました。村岡宏是先輩、河合利剣先輩、稲葉翔先輩には気軽に研究の相談にのっていました。数多くの先輩に協力していただき、研究を進めることができました。深く感謝いたします。

研究室の同期も公私に渡って心の支えになっていただきました。特に高本聰くんに

は研究の手法，解析方法においても相談し，有益なアドバイスをもらいました．深く感謝しています．

最後に，研究に理解を示し常に支えてくれた家族に感謝いたします．ありがとうございました．

## 参考文献

- [1] A. Bandura. Self-efficacy:toward a unifying theory of behavioral change. *Psychological review*, Vol. 84, No. 2, p. 191, 1977.
- [2] D. E. Berlyne. *Conflict, Arousal, and Curiosity*. McGraw-Hill Book Campany, 1960.
- [3] Lois Bly, 木本孝子, 中村勇. 写真で見る 乳児の運動発達 -生後 10 日から 12 ヶ月まで-. 協同医書出版社, 1998.
- [4] K. Doya. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, Vol. 10, No. 6, pp. 732–739, 2000.
- [5] A. G. Barto et al. Intrinsically motivated learning of hierarchical collections of skills. *IEEE International Conference on Development and Learning*, 2004.
- [6] A. Iriki et al. Coding of modified body schema during tool use by macaque post-central neurones. *NeuroReport*, Vol. 7, pp. 2325–2330, 1996.
- [7] A. Slater et al. Habituation in the newborn. *Infant Behavior and Development*, Vol. 7, No. 2, pp. 183–200, 1984.
- [8] A. Smith et al. Dopamine,prediction error and associative learning:a model-based account. *Network: Computation in Neural Systems*, Vol. 17, No. 1, pp. 61–84, 2006.
- [9] G. H. Mulliken et al. Forward estimation of movement state in posterior parietal cortex. *PANS*, Vol. 105, No. 24, pp. 8170–8177, 2008.
- [10] G. Taga et al. Analysis of general movements of infants towards understanding of developmental principle for motor control. *IEEE SMC '99 Conference*, Vol. 5, pp. 678–683, 1999.
- [11] H. Imamizu et al. Human cerebellar activity reflecting an aquired internal model of a novel tool. *Nature*, Vol. 403, pp. 192–195, 2000.
- [12] K. S. Narendra et al. Adaptation and learning using multiple models, switching and

- tuning. *IEEE Control Systems*, Vol. 15, No. 3, pp. 37–51, 1995.
- [13] K. Samezima et al. 複数の予測モデルを用いた強化学習による非線形制御. 電子情報通信学会論文誌, Vol. J84-D-II, No. 9, pp. 2092–2106, 2001.
- [14] M. Haruno et al. Mosaic model for sensorimotor learning and control. *Neural Computation*, Vol. 13, pp. 2201–2220, 2001.
- [15] M. Usher et al. The role of locus coeruleus in the regulation of cognitive performance. *Science*, Vol. 283, No. 22, pp. 549–554, 1999.
- [16] P. Oudeyer et al. Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, Vol. 11, pp. 265–286, 2007.
- [17] P. R. Montague et al. A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Neuroscience*, Vol. 16, No. 5, pp. 1936–1947, 1996.
- [18] R. Kawagoe et al. Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neuroscience*, Vol. 1, No. 5, pp. 411–416, 1998.
- [19] R.A. Jacobs et al. Adaptive mixtures of local experts. *Neural Computation*, Vol. 3, pp. 79–87, 1991.
- [20] Sarah-J. Blakemore et al. Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, Vol. 11, No. 5, pp. 551–559, 1999.
- [21] Sarah-J. Blakemore et al. Action prediction in the cerebellum and in the parietal lobe. *Experimental Brain Research*, Vol. 153, No. 2, pp. 239–245, 2003.
- [22] Sarah-J. Blakemore et al. Delusions of alien control in the normal brain. *Neuropsychologia*, Vol. 41, pp. 1058–1067, 2003.
- [23] W. Schultz et al. A neural substrate of prediction and reward. *Science*, Vol. 275, No. 5306, pp. 1593–1599, 1997.
- [24] 浅田 稔 al. 認知ロボティクスの目指すもの. 日本ロボット学会誌, Vol. 17, No. 1, pp. 1–5, 1999.
- [25] L. Festinger. *A theory of Cognitive Dissonance*. Stanford University Press, 1957.
- [26] C. Fiorillo. The uncertain nature of dopamine. *Molecular Psychiatry*, Vol. 9, No. 2, pp. 122–3, 2004.
- [27] S. Gallagher. Philosophical conceptions of the self:implications for cognitive science. *Trends in Cognitive Sciences*, Vol. 4, No. 1, pp. 14–21, 2000.
- [28] E. V. Holst. Relation between the central nervous system and peripheral organs. *British Journal of Animal Behavior*, Vol. 2, pp. 84–94, 1954.
- [29] M. S. Hook and P. W. Kalivas. Involvement of dopamine and excitatory amino acid transmission in novelty-induced motor activity. *The Journal of Pharmacology and*

- Experimental Therapeutics*, Vol. 269, No. 3, pp. 976–988, 1994.
- [30] C. L. Hull. Principles of behavior: An introduction to behavior theory. *New-York Appleton-Century-Croft*, 1943.
- [31] J. M. Hunt. Intrinsic motivation and its role in psychological development. *Nebraska Symposium on Motivation*, Vol. 13, pp. 189–282, 1965.
- [32] A. M. Leslie and S. Keeble. Do six-month-old infants perceive causality? *Cognition*, Vol. 25, No. 3, pp. 265–288, 2005.
- [33] U. Neisser. Five kinds of self-knowledge. *Philosophical Psychology*, Vol. 1, pp. 35–59, 1988.
- [34] R. D. Oades. The role of noradrenaline in tuning and dopamine in switching between signals in the cns. *Neuroscience and Biobehavioral Reviews*, Vol. 9, No. 2, pp. 261–82, 1985.
- [35] J. Panksepp. *Affective neuroscience:the Foundations of Human and Animal Emotions*. Oxford University Press, 1998.
- [36] H. F. R. Prechtl. *Continuity of Neural Functions from Prenatal to Postnatal Life*. Cambridge University Press, 1984.
- [37] H. F. R. Prechtl and B. Hopkins. Developmental transformations of spontaneous movements in early infancy. *Early Human Development*, Vol. 14, No. 3-4, pp. 233–238, 1986.
- [38] J. N. J. Reynolds and J. R. Wickens. A cellular mechanism of reward-related learning. *Nature*, Vol. 413, pp. 67–70, 2001.
- [39] Aldebaran Robotics. Nao datasheet h21 h25. <http://www.aldebaran-robotics.com>, 2011.
- [40] Aldebaran Robotics. Nao links v4.0. <http://www.aldebaran-robotics.com/documentation/nao/hardware/kinematics/nao-links-40.html>, 2011.
- [41] Aldebaran Robotics. Nao video camera. <http://www.aldebaran-robotics.com/documentation/nao/hardware/video.html>, 2011.
- [42] Aldebaran Robotics. Naoqi overview. <http://www.aldebaran-robotics.com>, 2011.
- [43] P. Rochat. Five levels of self-awareness as they unfold early in life. *Consciousness and Cognition*, Vol. 12, No. 2, pp. 717–731, 2003.
- [44] P. Rochat and S.J. Hespos. Differential rooting response by neonates: Evidence for an early sense of self. *Early Development and Parenting*, Vol. 6, pp. 105–112, 1997.
- [45] P. Rochat and T. Striano. Emerging self-exploration by 2-month-old infants. *Developmental Science*, Vol. 2, No. 2, pp. 206–218, 1999.

- [46] R. M. Ryan and E. L. Deci. Intrinsic and extrinsic motivations:classic definitions and new directions. *Contemporary Educational Psychology*, Vol. 25, No. 1, pp. 54–67, 2000.
- [47] J. Schmidhuber. Curious model-building control systems. *Neural Networks*, Vol. 2, pp. 1458–1463, 1991.
- [48] J. Schmidhuber. Curious model-building control systems. *In Proceeding International Joint Conference on Neural Networks*, Vol. 2, pp. 1458–1463, 1991.
- [49] W. Schultz. Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, Vol. 80, No. 1, pp. 1–27, 1998.
- [50] L. B. Smith and E. Thelen. Development as a dynamic system. *Trends in Cognitive Sciences*, Vol. 7, No. 8, pp. 343–348, 2003.
- [51] J. C. Stanley. Computer simulation of a model of habituation. *Nature*, Vol. 261, pp. 146–148, 1976.
- [52] R. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, Vol. 3, No. 1, pp. 9–44, 1988.
- [53] R. S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, Vol. 3, pp. 9–44, 1988.
- [54] R. S. Sutton and A. Barto. *Reinforcement Learning:An Introduction*. Cambridge,MA/MIT Press, 1998.
- [55] E. Thorndike. *Animal Intelligence:Experimental Studies*. MacMillan,New York,NY, 1911.
- [56] L. S. Vygotsky. *Mind in society: the development of higher psychological processes*. Harvard University Press, 1978.
- [57] H. Watanabe and G. Taga. General to specific development of movement patterns and memory for contingency between actions and events in young infants. *Infant Behavior and Development*, Vol. 29, No. 3, pp. 402–422, 2006.
- [58] J. S. Watson. Smiling, cooing and 'the game'. *Journal of Developmental Psychology*, Vol. 18, , 1972.
- [59] R. W. White. Motivation reconsidered:the concept of competence. *Psychological Review*, Vol. 66, No. 5, pp. 297–333, 1959.
- [60] D.M. Wolpert and M. Kawato. Multiple paired forward and inverse models for motor control. *Neural Networks*, Vol. 11, pp. 1317–1329, 1998.
- [61] ウーシャ・ゴスワミ. 子どもの認知発達. 新曜社, 2003.
- [62] ジャンピアジエ. 知能の誕生. ミネルヴァ書房, 1978.
- [63] 独立行政法人科学技術振興機構. 小脳に獲得される道具の内部モデル.

- 
- . [http://www.jst.go.jp/erato/project/kgd\\_P/kgd\\_P/kgd05.html](http://www.jst.go.jp/erato/project/kgd_P/kgd_P/kgd05.html)
- [64] 山内光哉. 発達心理学—上 [第2版] 周産・新生児・乳児・幼児・児童期. ナカニシヤ出版, 1998.
- [65] 山内宏太朗. 人間の発達を考える1 胎児から青年まで. 北樹出版, 1997.
- [66] 相馬幸子. ピアジェの感覚運動的知能. 北星論集, Vol. 25, pp. 267–299, 1988.
- [67] 銅谷賢治. 計算神経科学への招待：脳の学習機構の理解を目指して. サイエンス社, 2007.