OUKA
Osaka University Knowledge Archive

Title	情報量に基づく脚式移動ロボットのための観測行動戦略
Author(s)	光永, 法明
Citation	大阪大学, 2003, 博士論文
Version Type	VoR
URL	https://hdl.handle.net/11094/2355
rights	
Note	

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

情報量に基づく
脚式移動ロボットのための観測行動戦略

光永 法明

2002年12月

謝辞

本研究を進めていく上で、懇切丁寧な御指導、多くの助言、有益な議論をして頂いた大阪大学大学院工学研究科浅田稔教授に深い感謝の意を表します。

数々の議論、助言を頂いた大阪大学大学院工学研究科細田耕助教授に深い感謝の意を表します。

また貴重な討論、数々の意見を頂いた大阪大学大学院工学研究科浅田研究室で同席した学生各位に感謝の意を表します。

そして研究生活を支えてくれた両親に感謝します。

目 次

第1章 序論 **13**
1.1 移動ロボットの従来研究の流れ 13
1.2 状態空間構成に関する研究の流れ 14
1.3 本研究の目的と手法 15

第2章 従来研究 **17**
2.1 状態空間表現と状態推定法 17
2.1.1 カルマンフィルタを用いた自己位置推定と位置表現 18
2.1.2 マルコフ自己位置推定と位置表現 19
2.1.3 ニューラルネットによる状態表現 19
2.1.4 強化学習研究における状態空間構成 20
2.2 効率的観測に関する研究 21
2.2.1 観測計画 21
2.2.2 予測精度に基づく観測 22
2.2.3 エントロピーを減少させる観測 23
2.3 移動中の観測の利用 24
2.4 本研究の状態表現と観測戦略 26

第3章 情報量による観測対象選択 **29**
3.1 緒言 29
3.2 情報量による観測対象選択 29
3.2.1 環境設定 29
3.2.2 観測基準としての情報量 30
3.2.3 決定木，予測木の生成 31
3.2.4 確率の計算 32
3.2.5 行動決定 33
3.3 実験 34
3.3.1 実験環境 34
3.3.2 実験 1 36
3.3.3 実験 2 41

3.4 討論 43
3.5 結言 44

第4章 情報量による観測対象と注視窓の選択 45
4.1 緒言 45
4.2 自律的センサ空間分割の重要性 45
4.2.1 視覚センサの特性と注視窓 46
4.3 情報量に基づく効率的観測と行動決定 47
4.3.1 環境設定 47
4.3.2 観測による情報量 47
4.3.3 観測時間の考慮 48
4.3.4 行動決定木の生成 48
4.3.5 行動決定 49
4.3.6 履歴を考慮した行動決定木の再生成 51
4.4 実験 53
4.5 実験結果 55
4.6 結言 67

第5章 歩行ロボットのための情報量に基づく注視制御 69
5.1 緒言 69
5.2 歩行中の観測を考慮した注視制御・行動決定問題 70
5.2.1 必要となる画像補正 70
5.2.2 歩行中の観測を考慮した意思決定 71
5.3 情報量による注視選択と歩行への対応 71
5.3.1 行動決定 72
5.3.2 画像の補正法 73
5.4 実験 75
5.4.1 ロボットと実験環境 75
5.4.2 画像の補正実験 76
5.4.3 提案手法による行動決定の様子 80
5.5 観測行動決定閾値についての考察 81
5.6 結言 83

第6章 結言 85

付 録A 歩行の獲得 89
A.1 従来研究 90
A.2 ダイナミクスを生かした歩行獲得 90

A.2.1 運動モデルと歩行パラメータ 91
A.2.2 歩行パラメータの探索法 92
A.2.3 脚先接地センサによる歩行実現度 92
A.3 実験結果 92
A.3.1 ロボットの構造と脚先軌道 92
A.3.2 歩行パラメータの変動と歩行実現度, 移動速度 93
A.3.3 歩行実現度と移動速度 95

図目次

2.1 Action decision based on self-localization. 18
2.2 Action decision which does not rely on self-localization. 18
2.3 Recurrent nueral network for state estimation. 20

3.1 Example classificatoin tree by ID3. 32
3.2 An example of an action decision tree. 33
3.3 An example of a prediction tree for landmark A. 34
3.4 The SONY legged robot for RoboCup 99 SONY legged robot league. . . . 35
3.5 Experimental field (same as the one for RoboCup SONY legged robot league). Cross and circle marks are for the first experiment. 35
3.6 Quantization for landmarks and the ball. 36
3.7 A part of the action decision tree (Experiment 1). F, L, and R mean forward, left forward, and right forward respectively. 37
3.8 Robot movements with the action and decision trees. 39
3.9 Probability distribution in Experiment 1-1 (The gray level of each box indicates the probability, black 1 and white 0) 40
3.10 Probability distribution in Experiment 1-2. 40
3.11 Probability distribution in Experiment 1-3. 41

4.1 Quantization for landmarks and the ball. 46
4.2 The action decision tree constructed from example training data 51
4.3 Calculation example of action probabilities 52
4.4 An example training sequence and re-constructed action decision tree . . . 54
4.5 A robot for the RoboCup SONY legged robot league (a) and prepared moving actions (b). 55
4.6 The experimental environment (the same configuration for RoboCup SONY legged robot league) 56
4.7 Generated attention windows with pre-quantized sensor values 57
4.8 Generated attention windows with quantization by information gain 57
4.9 Generated attention windows with quantization by information gain per time 58

4.10 Generated attention windows of the re-constructed tree. Quantization was done by information gain per time. 58
4.11 Generated attention windows of the re-constructed tree. Quantization was done by information gain per time. Action decisions started from middle of the training sequence were considered 59
4.12 An example action decision sequence by the tree constructed by information gain. The probabilities to reach each leaves (left) and the action probabilites (right) are shown. The intensity of each small box indicates the probability, white means zero and black means one. 61
4.13 An example action decision sequence by the tree constructed by information gain per time. The probabilities to reach each leaves (left) and the action probabilites (right) are shown. 62
4.14 An example action decision sequence by the re-constructed tree with action state. The original tree was constructed by information gain per time. The probabilities to reach each leaves (left), the action state probabilities (center), and the action probabilities (right) are shown. 62
4.15 Another example action decision sequence by the tree constructed by information gain per time. We used 53 training data for this tree. The probabilities to reach each leaves (left) and the action probabilites (right) are shown. 63
4.16 A part of the action decision tree constructed by the proposed method . . 64
4.17 Changes of attention windows along with the increase of data(1/2) 65
4.18 Change of attention windows along with the increase of data(2/2) 66

5.1 An example action decision tree. 72
5.2 A time sequence example. At time A, it starts walking. A, B, C,... are the beginning of a walking period. 73
5.3 Experimental field (half of the RoboCup 2002 SONY legged robot league). Robot started from three positions on the center line. 75
5.4 The images of the camera while the robot is moving forward. The images are taken every 80[ms]. 76
5.5 The $\triangle y$ of forward motion. 77
5.6 The $\hat{\boldsymbol{x}}_S$, and $\boldsymbol{\sigma}_S$ of the forward motion watching at front direction. 77
5.7 The $\hat{\boldsymbol{x}}_S$, and $\boldsymbol{\sigma}_S$ of the right forward motion watching at front direction. . 78
5.8 The $\hat{\boldsymbol{x}}_S$, and $\boldsymbol{\sigma}_S$ of the left forward motion watching at front direction. . . 78
5.9 Image compensation to the x axis by proposed method. 79
5.10 Image compensation to the y axis by proposed method. 79

5.11 Generated attention windows by the proposed method. 80
5.12 Changes of expected information gain, the highest action probability, and actions. One of which has the highest action probability, and the other is the one taken by the robot. The starting point of robot was the center of the field. The action probability threshold was 0.4 and the information gain threshold was not used. 82
5.13 Changes of expected information gain, the highest action probability, and actions. One of which has the highest action probability, and the other is the one taken by the robot. The starting point of robot was the center of the field. The action probability threshold was 0.6 and the information gain threshold was not used. 83
5.14 Changes of expected information gain, the highest action probability, and actions. One of which has the highest action probability, and the other is the one taken by the robot. The starting point of robot was the center of the field. The action probability threshold was 0.4 and the information gain threshold was 0.4. 84
5.15 The values of expected information gain and maximum action probability observed in experiments. 84

A.1 The side and front view of the legged robot 93
A.2 Landing point of leg versus walk speed and realized degree of walk. 94
A.3 Angle of supporting line versus walk speed and realized degree of walk. . . 94
A.4 Realized degree of walk by foot contact switches versus average speed . . . 95

表目次

2.1 Comparison of observation strategies for a mobile robot 24
2.2 Comparison of methods for observation during its locomotion 26

3.1 Example training data of ID3. 31
3.2 An example of teaching data. 32
3.3 Depth of the prediction trees(Experiment 1). 38
3.4 The order of information for the action decision tree (Experiment 1). . . . 38
3.5 The order of information for prediction trees (Experiment 1). 38
3.6 The number of needed observation (Experiment 1). 41
3.7 Depth and size of the action decision tree (Experiment 2). 42
3.8 Depth and size of the prediction trees (Experiment 2). 42
3.9 The order of information for the action decision tree(Experiment 2). 42
3.10 The order of information for prediction trees (Experiment 2). 43

4.1 Comparison of the number of leaves between the previous method and the C4.5 46
4.2 Example training data 49
4.3 Information gain and information gain per time calculated from example training data (Info., Info./time and Lm indicate information gain, information gain per time and landmark respectively). 50
4.4 Comparison of the size of trees, expected number of gaze directions, and time to make a decision. 59
4.5 Comparison of average number of gaze directions, and time to make a decision in experiments. 60
4.6 Comparison of the size of trees, expected number of gaze directions, time to make a decision, and average attention window sizes. 63

第1章　序論

ロボットには人間には困難であったり面倒な作業, また人間が近づけない場所における作業を行なうことが期待されている．すなわち，ロボットは様々な場所へ移動し，場所に応じた行動をとることが期待される．そのためロボットは持っているセンサから，適切な行動決定をしなければならない．

移動ロボットには視覚センサが搭載されることが多いが，視覚センサにより得られる情報は膨大であり，効率よく行動決定に必要な情報を抽出する必要がある．注視によりシーンの理解など視覚の問題が解きやすくなることは多く報告されており，タスクに応じた選択的注視法が求められている [1]．注視には，ある視覚目標を見続ける注視制御，環境地図の獲得など計測を簡単にするための注視，ロボットの行動決定のための注視などの問題がある．また選択する視覚情報としては，画像特徴，視覚の解像度，観測方向，観測対象物などが挙げられる．移動ロボットに搭載される視覚センサ (カメラ) の多くは知覚可能範囲 (視野角) が限られており，環境中からより多くの情報を得るためには，能動的にセンサの検出方向を変化させる能動視覚が必要となる．したがって，移動ロボットでは能動視覚を考慮した行動決定のための注視，観測戦略が重要となる．

本章では，自己位置/環境表現を明に与える移動ロボットに関する従来研究の流れと，自己の状態表現をロボット自身が獲得することを目指す研究の流れを紹介し，それぞれの観測戦略について概観する．そして従来研究に対する本研究の立場を明らかにし，本研究の目的と本論文の内容と構成について述べる．

1.1　移動ロボットの従来研究の流れ

移動ロボットにおいては，ロボット自身の位置に応じてタスクを実行する上で必要とされる行動が変化することから，1) まず自己位置を決定し，2) 必要に応じてその他の物体を認識し，3) 行動の決定や計画のやり直しを行うアプローチが多くとられてきた．地図と自己位置の表現には大きく分けて二つの方法が用いられている．一つはトポロジカルな表現であり，環境中の場所をドアや通路，部屋，建物などシンボルで表し，シンボル間を左右や東西南北などの幾何学的位置関係子などでつないだものを地図としている．この表現を用いると地図がコンパクトではあるが正確な位置の表現には向かない．もう一つは定量幾何学的な表現であり，環境中の位置を幾何学と数値で表したものである．この表現を

用いると地図や計算コストは高くなるが，位置を正確に表すことができることから，多くの移動ロボットに用いられている．

自己位置の認識のためには環境の地図を用意する必要があるため，地図の獲得に関する研究 [2, 3, 4, 5, 6, 7] が行われている．またセンサ情報の統合と地図の参照による自己位置同定に関する研究がさかんに行われている．移動ロボット研究においては障害物を回避しながら目的地へ到達するナビゲーションの問題に焦点があてられることも多い [8, 9]．

観測地点の計画や，ランドマークの配置計画についての研究も行われているが，ナビゲーションに特化した戦略が多い．また，こういった研究の多くは車輪式移動ロボットを前提としているため，脚式移動ロボットの場合には向かない面もある．これらの研究ではデッドレコニングのモデルを利用しているが，一般に場所に依存せず，一定の行動には一定の移動量を対応させている．しかし歩行による移動の場合，移動距離がすべりなどのため歩行面へ依存して大きく変化し，移動量のモデルは場所に応じたものが必要となり，あらかじめ移動量のモデルを用意することは困難な場合がある．

1.2 状態空間構成に関する研究の流れ

場所の区別は必要であっても，必ずしも通常の幾何学的位置の表現形態をとる必要はない場合が考えられる．視覚を持つ移動ロボットを考えた場合，視覚情報から幾何学的位置を再構成するには，計算コストだけでなく，ロボット自身と環境についてのモデルやパラメータといった知識が必要となる．幾何学的表現を用いずに，視覚情報をそのまま用いることが可能であれば，そういった知識は必要ない．また位置の計算のための観測とそれ以外の意思決定に必要な観測の区別や，視覚センサとそれ以外のセンサの区別なしに扱うことも可能となる．しかし，視覚センサにより得られる情報は膨大であり，そのまま扱うことは難しく，なんらかの抽象化が必要となることが多い．

強化学習を用いたロボットの学習においては，学習時間が状態空間の大きさに対して指数関数的に大きくなる [10] ことから，センサ空間をそのまま状態空間として用いることが出来ないことが多い．そのため自律的センサ空間の分割による状態空間構成法がいくつか提案されている．しかし，これらの手法では観測にかかる時間や，能動的視覚は考慮されていない．

また，状態の表現としてセンサ情報をそのまま利用し，センサ情報とコンテキストベクトルからなるリカレントニューラルネットワークを用いて，次のセンサ情報 (場所) を予測することにより，ある種の自己位置表現を獲得する研究が行われている [11, 12, 13, 14]．そして予測精度に応じて能動的に観測方向を変える実験などが行われている [15]．しかし，状態表現獲得の安定性や，センサ表現への工夫が必要など問題が残っており，観測戦略についてはさらなる研究が必要である．

1.3 本研究の目的と手法

本研究の目的は，視覚センサを持つ脚式移動ロボットの行動決定のための効率のよい観測戦略を提案することである．観測戦略の効率は行動決定までの観測時間の短さであるとし，明示的な自己位置同定は行わない．観測の行動決定への寄与を情報量により比較し，寄与の大きい観測から実行することにより，静止観測時の効率的な観測戦略を実現する．さらに行動決定の効率の向上を図るため，移動中の観測を行動決定に利用できるよう手法を拡張する．

まず，ランドマークの観測により得られる行動に関する情報量により行動決定木を生成し，決定木を用いてランドマークの観測順序と行動を決定する手法を提案する．「行動に関して得られる情報量が大きい」とは，行動に関する曖昧さが減少する，すなわち行動決定に向くということを表す．そこで，トレーニングデータから情報量を計算することにより，情報量の大きいランドマークから順に観測し，行動決定を行う決定木と観測予測を行う予測木を生成する．行動決定の際には予測と，行動決定木にしたがって，順にランドマークを観測することにより，観測時間の短い行動決定を実現できる．

次に，観測に要する時間と，視覚センサの特性を考慮したセンサ空間の自律的な構成により状態を生成する，観測時間のより短い行動決定手法を提案する．観測に要する時間は一定ではなく，能動視覚において視覚センサの向きを変更するには比較的長い時間がかかる．そこで情報量ではなく，単位時間当りに得られる情報量を用いることで観測時間の短縮をはかる．また視野角の限られた視覚センサを考慮するため，観測戦略には，どのランドマークを観測対象とするかだけでなく，どの方向を観測するかの記述があることが望ましい．そこで観測をランドマークが注視窓と呼ぶある範囲に見られるか否かを知ることと定義し，その観測についての単位時間当りの情報量を用いて行動決定木を生成することを提案する．これにより，より観測時間の短い行動決定を行うことができる．

そして，自律的センサ空間構成により得られたセンサ空間を用いて，行動決定木を再生成することにより，履歴を考慮した行動決定木を生成する．行動決定木の生成と同時に得られる自律的センサ空間は，行動が同じであるセンサ空間となっており，ある種の状態と見なすことができる．一般に状況は連続して変化し，ロボットがとるべき行動を決定する際に過去の状態が分かることで行動決定に必要な観測が減少することが期待できる．そこで構成されたセンサ空間を用いて一時刻前の状態を表し，その状態を利用して行動決定木を生成することで更なる観測時間の減少をはかる．

さらに，歩行によるセンサの揺れの補正と，異なる観測位置でのセンサ値の統合を行うため，歩行中の画像列から補正値を計算し，行動決定時に利用する．補正値の計算には画像列のみを用い，他のセンサは用いない．補正値を用いて，静止時にある観測が得られる確率を計算することで，複数観測の統合と揺れの補正を実現する．また，行動決定の曖昧さを表す指標として，観測により得られることが期待される情報量を用いることで，よりよい行動決定の実現をはかる．

以下に本論文の構成を示す．第2章では，観測戦略と密接な関係のある状態空間表現と状態推定法の従来研究，観測戦略の従来研究，移動中の観測の利用についての従来研究について紹介すると共に，問題点を指摘し，本論文の立場を明らかにする．第3章では，ランドマークの観測により得られる行動に関する情報量により行動決定木を生成し，決定木を用いてランドマークの観測順序と行動を決定する手法を提案する．まず，曖昧さを表す量エントロピーと情報量の定義を示し，情報量を用いた行動決定木の生成法と，決定木を用いた観測戦略と行動決定について述べる．そして小型四脚ロボットを用いた実験により，手法の有効性を示す．第4章では，観測に要する時間と，視覚センサの特性を考慮した自律的なセンサ空間の構成と履歴の考慮により，観測時間のより短い行動決定手法を提案する．まず，自律的センサ空間構成の重要性を実験データから示す．次に視覚センサの特性から注視窓を提案し，自律的センサ空間構成と行動決定木の生成を同時に行う手法と，行動決定木を利用した，観測と行動決定法について述べる．そして，構成された自律的センサ空間を用いて行動決定木を再生成することにより，履歴を考慮した決定木を生成する手法を提案する．最後に実験により，実際に観測時間が短縮されていることを示す．第5章では，歩行によるセンサの揺れの補正と，異なる観測位置でのセンサ値の統合を行うため，歩行中の画像列から補正値を計算し，行動決定時に利用することを提案する．補正値を用いて，静止時にある観測が得られる確率を計算することで，複数観測の統合と揺れの補正を実現する．また，これから観測により得られると期待される情報量を行動決定の曖昧さの指標として用い，行動確率あるいは情報量期待値が閾値を満足しない場合には，移動せずに観測を行うことを提案する．実験により，提案手法の有効性を検証する．最後に第6章で，結論と将来の研究課題について述べる．

第2章　従来研究

本章では，観測戦略と密接な関係のある状態空間表現と状態推定法についての従来研究を紹介し，次に観測戦略と移動中の観測についての従来研究を紹介すると共に問題点を指摘する．最後に本論文で提案する利用する状態空間表現と観測戦略の従来研究に対する立場を示す．

2.1　状態空間表現と状態推定法

移動ロボットの内部表現には何らかの形で自己位置の表現が必要となることがほとんどである．またロボットの持つセンサの制約のため，自己位置の推定には複数のセンサ値，あるいはセンサ値の履歴の統合が必要となることが多い．位置の表現形式としては，

- 相対的な位置関係を状態遷移上で記述する位相幾何学的 (トポロジカルな) 位置表現
- 直交座標系を用いた定量幾何学的位置表現
- センサ空間の直積
- センサ値の履歴をリカレントニューラルネットワークで統合した表現
- センサ空間を自律的に構成した表現

などが挙げられる．自己位置の表現を明示的に行う手法での観測戦略と行動決定の流れは，**Fig.2.1** のように，まず自己位置同定のための観測行動を行い，自己位置が決定されてからゴールへ向かう移動行動決定を行う．利点としては，観測戦略と行動決定が分離しているため，自己位置表現が同じであれば異なるロボットでも同じ行動決定部分を再利用できる可能性が高いという，汎用性が挙げられる．しかし，センサ値から自己位置の獲得と，自己位置から行動への変換の2段階が必要であり，全体としての行動決定時間が長くなる．一方，本研究を含む自己位置表現を明に行わない手法では，観測戦略と行動決定の流れは，**Fig.2.2** のように，自己位置表現をとらず，直接観測行動とゴールへ向かう移動行動を決定する．観測戦略は自己位置にとらわれず，行動決定に必要なだけの観測を行う．そのため，より素早い行動決定が行える点が利点となる．

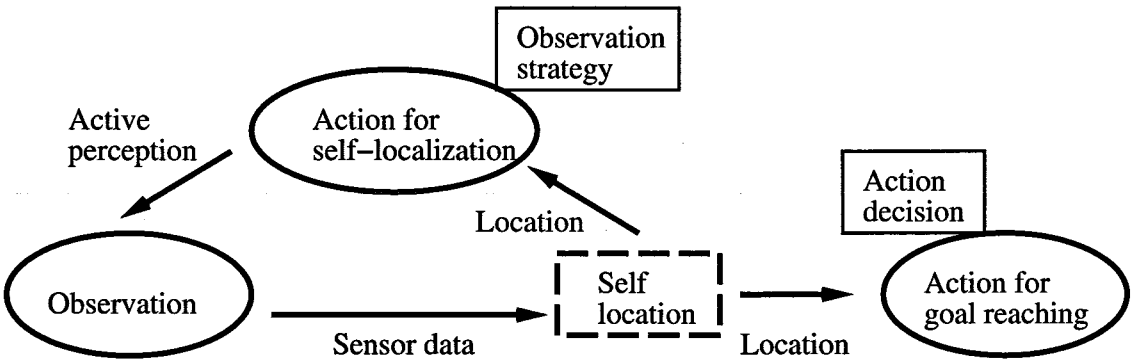


Fig.2.1 Action decision based on self-localization.

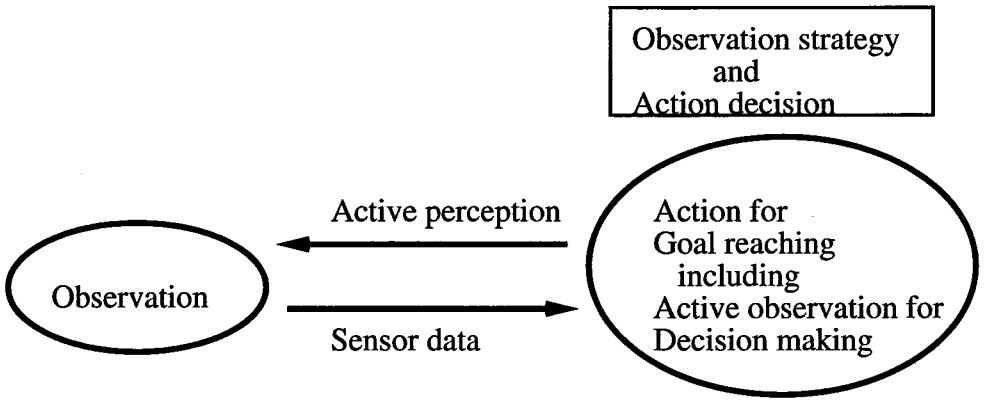


Fig.2.2 Action decision which does not rely on self-localization.

以下では定量幾何学的位置表現におけるカルマンフィルタによる観測統合，位相・定量幾何学的位置表現におけるベイズ推定による観測統合，リカレントニューラルネットワークによる表現の獲得，センサ空間の自律的構成についての従来研究を紹介し，それぞれの問題点を指摘する．

2.1.1 カルマンフィルタを用いた自己位置推定と位置表現

カルマンフィルタは，3次元視覚の観測統合 [16] や，移動ロボットの観測統合と自己位置の更新 [17] への利用が提案され，自己位置同定の際の複数の仮説の更新 [18, 19] や，自己位置同定を伴う地図の生成 [20, 21, 22, 23] などに広く用いられている．

ロボットの自己位置は，直交座標上で平均と分散により表す．内界センサ (移動距離) により更新された自己位置と，外界センサによるランドマークの観測と地図から計算した

自己位置をカルマンフィルタにより統合し，自己位置の推定値(平均と分散)を更新する．一般に3次元再構成する視覚処理は計算量が多く，時間がかかることから，予測による視覚処理の軽減[24]や，視覚処理と移動を並行して行ない，視覚処理が終了した時点で推定位置を更新する方法[25, 26]が提案されている．

一つの自己位置モデルは一つの平均と分散しか表すことができない．そのため自己位置が不明な状態から位置を同定するためには，複数の自己位置仮説を用意し，それぞれを更新すること[18, 19]が行われている．また自己位置の更新に必要な計算量は少ないが，センサ情報から自己位置を計算する処理が必ず必要であり，処理は3次元再構成や地図との照合などを含む複雑なものとなる問題がある．

2.1.2 マルコフ自己位置推定と位置表現

マルコフ自己位置推定では，自己位置と方向の遷移がマルコフ過程であると仮定し，各自己位置と方向の候補に確率を割り当てる．移動や観測した際にベイズ推定により，位置と方向の確率を更新する．この手法は位相幾何学的表現にも定量幾何学的表現にも用いることができる．大きく分けてトポロジカルな地図上であるノードの示す位置である確率による表現[27, 28, 29]と，自己位置と方向の直積空間を格子状に分割し，現在の位置と方向が格子で表される範囲内である確率により表現する手法[30, 31]の2種の手法が提案されている．複数の自己位置候補を確率で表現していることから，カルマンフィルタにより自己位置を更新する手法と異なり，自己位置が不明な状態からも観測を繰り返すことにより，自己位置を推定することが可能である．また格子により自己位置を表現した場合には，自己位置候補の広がる空間を表す形状に制限がないことから，完全に3次元再構成できないセンサ情報も利用することができる．またノイズにも強いことが実験的に示されている[32]．一方，確率マップ表現をとるため，マップの更新に必要な計算量は大きい．

トポロジカルな表現は，シンボリックな表現であり，コンパクトであるが詳細な位置表現とはいえず，適切なシンボル表現を用意する必要がある．一方，格子により表現する場合には，格子の切り方により詳細な表現が可能である．しかし，記憶量と計算量が大きくなるため記憶量や計算量の圧縮について研究されている[31]．また格子全体の計算をするのではなく，代表点のみに計算を抑えたモンテカルロ位置同定(Montecalro Localization: MCL)も提案されている[2, 33, 34, 35, 36]．

2.1.3 ニューラルネットによる状態表現

現在のセンサ情報のみからは自己位置に相当する状態の特定ができず，センサ値の履歴が必要となる場合について，センサ値の予測を行うリカレントニューラルネットワークを構築し，その出力を利用することで状態の特定を行う研究がある[11, 12, 13, 14]．Tani [11]は，**Fig.2.3**に示すリカレントニューラルネットワークを用いている．ネットワーク

の入力は，現在のセンサ値，行動，一時刻前のコンテキストベクトルである．出力は1時刻後のセンサ値とコンテキストベクトルである．ネットワークが1時刻後のセンサ値を正しく予測するように教師付き学習を行う．正しくセンサ値が予測できるならば，状態を特定できていると考えられるから，現在のセンサ値とコンテキストベクトルを組としたものを状態として用いる．

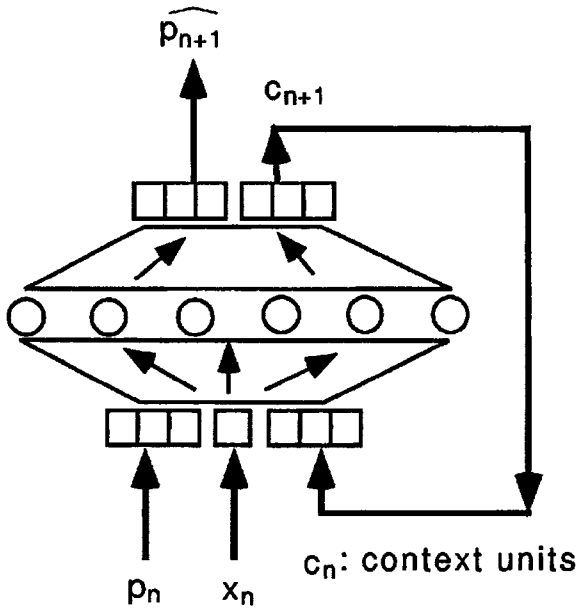


Fig.2.3 Recurrent nueral network for state estimation.

こういった手法の問題点としては，必要となるリカレントニューラルネットワークの複雑さが不明である，学習に時間がかかる，新たに見分けるべき状態が増えたときの対応が難しい，ある程度履歴がネットワークに蓄えられるまでコンテキストベクトルが正しくないなどが挙げられる．

2.1.4 強化学習研究における状態空間構成

強化学習の分野では，学習時間が状態空間の大きさに対して指数関数的に大きくなる[10]．また学習に用いる状態空間によって問題の解きやすさが大きく変わり，学習時間に大きな影響を与える．一方，学習するロボットにはできるだけ事前に知識を与えたくない．例えば視覚を持つ移動ロボットを考えた場合，視覚情報から幾何学的位置を再構成するには，計算コストだけでなく，ロボット自身と環境についてのモデルやパラメータといった知識が必要となる．そこでセンサー空間からロボットが自律的に状態空間を構成する手法が提案されている[37, 38, 39, 40, 41, 42, 43, 44]．最終ゴールからの近さや，同一行動に

より得られる報酬が同じであることなどに注目して，センサー空間に部分空間を構成し状態とする．部分空間の構成は，センサ値が離散である場合には論理和をとる方法などが，センサ値が連続である場合には，センサー軸を2分する方法や，センサ空間上での超平面となる線形判別関数による2分する方法，部分空間を代表点を中心とした超楕円とする方法などが提案されている．しかし，観測にかかる時間や，能動的視覚システムは考慮されていない．注視の粗密制御が結果として得られことが示されている [40] が，用いたセンサは周囲の観測に能動制御を必要としない全方位視覚センサであった．

2.2 効率的観測に関する研究

効率的な観測を実現するために提案されている手法は大きく分けて，

- 観測計画，学習による行動政策の獲得
- 予測精度に基づく観測切り替え
- エントロピーに基づく観測選択

に分けられる．以下では，それぞれについての従来研究を紹介し，問題点を指摘していく．

2.2.1 観測計画

ステレオカメラを用いて左右の視差から奥行きの計算を行う移動ロボットでは，特に画像処理に時間がかかるため，ナビゲーションに支障が出ない程度に観測を減らすための観測点の計画が提案されている [45, 46]．これらの手法では，デッドレコニングを用いて移動しているとき，誤差の蓄積により，壁への衝突が起きないよう，ランドマークによる位置を補正する観測点を計画する．デッドレコニングとランドマークの観測の統合は，カルマンフィルタを用いた推定位置の更新式の形で定式化しており，推定誤差により壁への衝突などが起きる前にランドマークを観測する．この定式化では，観測の基準が位置誤差の大きさとなっており，ナビゲーションのための観測計画である．そのため，位置に関係しない他の行動決定のための観測を統一的に扱うことはできない．また注視選択は考慮されておらず，ランドマークは，その観測から位置を計算できるものに限られる．

同様の考え方でランドマークの配置計画も提案されている [47]．位置誤差が大きくなり衝突を起す前に，位置の推定誤差を減少させるランドマークが観測できるよう配置を行う．また，一定時間間隔でランドマークの観測により自己位置推定をする場合に，自己位置の推定誤差の最も少なくなると期待される方向を観測することが提案されている [48]．視覚の不確かさがある場合には，探索木の分岐が増加するため，プランニングコストが無視できなくなる．そこでプランニングコストを考慮した移動計画法が提案されている

[49]. 探索木の分岐の増加とプランニングコストの増加の関係を予め求めておき，適切なコストでプランニングを打ち切る．いずれもナビゲーションのための計画となっている．

Jensfelt et al.[50] は，観測のコストと，壁への衝突などの位置誤差に伴うコストの和を最小化する観測を提案している．移動経路や移動に伴う自己位置の更新と蓄積誤差のモデルは与えられることを仮定している．コスト関数を定義するため動的計画法 [51](Dynamic Programming: DP) の手法を用いて，コストに関して最適な観測を計画することができる．コストを位置以外の情報の取得に関するものを定義することで，センサを統一して扱うことができる．しかし，コストの定義に明確な指針はなく，問題に応じて定義する必要がある．

Bayes 推定を用いた移動と注視をともなう観測計画が提案されている [52] が，交叉点の識別を対象として，識別までの時間が最小となるよう計画を立てるのみであり，識別にともなう移動がナビゲーションや，それ以外のタスクへ与える影響は考慮されていない．また組み合わせの増大によるプランニングコストの問題は無視できない．

自己位置をトポロジカルマップ上で確率的に表現し，行動と観測による確率遷移を部分観測マルコフ過程 (Partially Observable Malkov Dicision Process: POMDP) として扱い最適政策を求めることが提案されている．理論的には観測まで含めて最適な移動計画が可能だが，位置表現としてはコンパクトなトポロジカル表現を用いた場合でも状態空間が大きすぎ，実際には最適政策を求めることは極めて困難であり，実現可能な方法が模索されている段階である [29, 53, 54, 55, 56].

Whitehead と Ballard[57] は，行動決定のための観測対象選択を含む強化学習法を提案している．観測対象選択により状態量圧縮を行い学習の効率化をはかっている．一つの観測対象が全体の状態を代表できると仮定することで，学習を実現している．しかし，一般にはそのような仮定は困難で，部分観測問題が生じるため学習できない．また連続量の取り扱いや，観測の統合などは考慮されていない．

McCallum[37] は，履歴を考慮した強化学習の状態空間構成法を提案し，提案手法を適切な車線変更を獲得するタスクに適用し，能動的観測行動を含む政策が学習により得られたと報告している．しかし，車線変更の前に必ず能動的観測を行わなければならない制約があった．一般には能動的観測が行動を制約するとは限らず，この手法が能動的な観測に有効であるかは疑問がある．また観測の効率については言及されていない．

2.2.2 予測精度に基づく観測

物理的制約や計算資源の問題などにより任意の時刻に最新の値を得ることの出来ないセンサ値について予測を行い，補間することが考えられる．ナビゲーションにおける位置の推定誤差が大きくなる前に観測により修正するのと同様に，センサ値の予測誤差が問題となるほど大きくならないよう観測することが提案されている [58, 15, 59]. Davison and Murray[58] は移動ロボットが自己位置を計算するために予測精度に応じて注視対象を選

択することを提案している.

Tani et al. [15] は，部分視覚を持つ移動ロボットが視覚目標追従を行いながら，壁に沿った移動を行う実験で，予測精度に応じて視覚目標と壁との間で観測対象を切り替える実験を行っている．この実験では場所によらず，ロボットが必要とする情報は壁と視覚目標物の見え方であり，意思決定に必要な情報は一定である．そのため，状況に応じて観測を効率的に行う手法とはなってない.

鈴木ら [59] は，能動的に観測方向を変えられるソナーセンサをもつロボットにおいて，予測誤差と情報の重要度の重み付き和による観測方向の決定を提案している．しかし，情報の重要度と，重みは，タスクに応じて定義する必要があるが，それぞれの適切な定義方法は提案されていない．また状況に応じて，必要なセンサと精度は異なると考えられるが，そういった問題は扱うことは困難である.

2.2.3 エントロピーを減少させる観測

確率で表された事象の曖昧さを表す量にエントロピーがある．エントロピーは事象 i の生起確率を p_i で表したとき，

$$H = \sum_i p_i \log p_i \tag{2.1}$$

で表される．観測前と観測後のエントロピーの差は，観測による曖昧さの減少量を表し，観測の情報量と呼ばれる．観測に要するコストが一定であるとき，このエントロピーを減少させる (情報量の大きい) 観測が，最も効率のよい観測となることから，物体の認識 [60]，素材の特定 [61]，状態量推定 [62]，自己位置同定 [63] などに用いられている.

阪口 [62] は，状態推定を目的として，エントロピーによる観測の逐次計画を提案している．複数の状態推定モデルを持つ場合に，使用する状態量推定モデルの選択と観測対象選択を情報量を元に行う．また状態量が連続値で確率分布がガウス分布に従う場合のカルマンフィルタを使った情報量の逐次計算の定式化を行い，計算量の減少をはかっている．提案手法を，カメラによる対象物の追従タスクにおいて，対象の移動モデル (直線・回転運動など) とカメラの焦点距離，方向の選択実験を行っている.

Burgard et al. [63] は，ナビゲーションの問題にエントロピーを用いている．自己位置と方向の直積空間を格子に切り，現在の位置と方向が格子で表した自己位置表現の元で，自己位置の曖昧さをエントロピーで測っている．彼らの想定したロボットでは，観測は必ず移動を伴うが，情報量の大きい移動と，目的地へ近づく移動は異なる場合があることから，移動行動 a をとった場合のエントロピー期待値と，移動行動 a の価値 $v(a)$ (DP の手法を用いて計算する) の重み付き和を最も小さくする行動 a^*,

$$a^* = \arg\min_a(E_a[H]) \tag{2.2}$$

をとることを提案している.

Table 2.1 Comparison of observation strategies for a mobile robot

			base of observation criterion			
	planning / learning	reactive	position error	sensor estimation error	cost function	entropy
文ら [45]	○		○			
Kidono et al.[46]	○		○			
城殿ら [48]	○		○			
三浦, 白井 [49]	○				○	
Jensfelt et al.[50]	○				○	
McCallum[37]	○				○	
Davison and Murray[58]		○		○		
Tani et al.[15]		○		○		
鈴木ら [59]		○			○	
Burgard et al.[63]		○				○ (position)
本研究	○	○				○ (action)

エントロピーを用いることにより，物体認識や状態，自己位置の推定における効率のよい観測 (状況に応じた観測選択) が同様に選択できる．しかし，これらの手法は，物体の認識や状態，自己位置の推定を目的としており，それ以降のタスクは考慮されていない．Burgard et al. の場合もナビゲーションしか考慮されていない．すなわち，一般に推定量に許される曖昧さは，状況に応じて変化するが，これらの手法では一つのタスクの実行中は一定であると仮定しているのが問題である．

Table 2.1 に，移動ロボットの観測戦略に関する従来研究と本研究の比較を示す．全体的な戦略 (計画あるいは事前の学習: planning / learning，反射的行動: reactive)，観測をする規範の違い (位置誤差: postion error，センサの予測誤差: sensor estimation error，コスト関数: cost function，位置に関するエントロピー: entropy(position)，行動に関するエントロピー: entropy(action)) で特徴づけられる．

2.3 移動中の観測の利用

移動ロボットが効率的に行動決定するためには，静止時での観測だけではなく，移動中の観測も利用できる必要がある．移動中の観測の利用と行動決定に関しては，

1) 移動中には静止時と比較してセンサに揺れなどの外乱が加わる，

2) 視覚情報の処理に比較的長い時間かかり，処理結果が利用可能になったときには無視できないほど移動している場合や，複数方向の観測 (首振り) の統合の際に観測間で移動が無視できない場合がある，

3) 観測の曖昧さから行動決定に曖昧さが残っている場合には，衝突などの問題を避けるため，移動速度を低下あるいは，静止し観測する必要がある，

等の問題がある．

視覚を持つ車輪式ロボットでは 1) については，考慮する必要があまりない．2) について Kosaka et al.[25]，前山ら [26] は，自己位置計算のためのデッドレコニングと視覚センサの統合において，視覚センサの処理による遅れを考慮した遡及的統合法を提案している．文ら [45] は，移動の不確かさを考慮した観測点計画を提案しているが，2) については遡及的統合法を，3) の問題については必要に応じて移動速度を低下させることで対処している．これらの遡及的統合法は，3 次元再構成による自己位置表現を基に行動決定する場合にしか利用できない．Fukase et al.[64] は，脚式ロボットのサッカータスクにおいて，状況に応じてランドマークの観測か、ボールを注視しながら移動するかを選択するため，動的計画法の手法を用いた計算を行うことを提案している．状態遷移を部分観測マルコフ決定過程として表し，移動行動と観測行動の行動価値を計算することで，2),3) の問題を解いている．しかし，観測方法は二者択一であり観測戦略とはいえない．また歩行を移動誤差の大きな移動とみることで，歩行に伴う揺れの問題を取り扱っていない．

揺れの少ないゆっくりとした歩行の場合を除き，脚式ロボットでは多くの場合に揺れは無視できない．宮下ら [65] は，視覚を持つ脚式ロボットが視覚目標物が同じに見えるようにロボットの胴体を揺動制御するとともに，転倒回避をロボットに用意することで，反射的静歩行が発現することを示した．彼らの実験では，目標が一定の見え方となるよう常に制御し，静歩行することで揺れの問題を非明示的に回避している．

受動的あるいは能動的にセンサの向きを制御することで揺れを吸収することも考えられるが，制御周期の短いフィードバック制御あるいは特別な機構が必要となる．一方，特別な制御機構は用いずに，得られたセンサ値を補正する手法も考えられる．岡田ら [66] は，歩行に伴う揺れの対策として，傾斜角センサにより胴体姿勢を計測し画像情報を補正することや，加速度センサにより遊脚着地時など振動の大きいときの画像を利用しないなどの工夫をしている．しかし両センサが利用できるのは，比較的揺れの小さい歩行の場合である．原ら [67] は，4 脚ロボットによる協調搬送の研究で，エンドエフェクタに取りつけた力センサの歩行による揺れを軽減するため，揺れモデルを用意しセンサ値を補正することを提案している．しかし，これらの研究では，注視制御や 2),3) の問題は考慮されていない．

Table 2.2 に，移動中の観測への対応の従来研究と本研究の比較を示す．移動形態 (車輪式: wheeled，脚式: legged) の違い，歩行による外乱 (disturbance by walking)，視覚情報処理の時間や観測間での移動による遅れ (observation delay) の考慮の有無，観測の曖昧さ

Table 2.2 Comparison of methods for observation during its locomotion

	wheeled / legged	disturbance by walking	observation delay	handling of ambiguity	observation strategy
Kosaka et al.[25]	W		○		
前山ら [26]	W		○		
文ら [45]	W		○	○	○
Fukase et al.[64]	L		○	○	
宮下ら [65]	L	△			
岡田ら [66]	L	○			
原ら [67]	L	○			
本研究 (第5章)	L	○	○	○	○

への考慮 (handling of ambiguity) の有無，観測戦略 (observation strategy) の有無で特徴づけられる．

2.4 本研究の状態表現と観測戦略

本研究においては，様々な事前知識が必要となる数値的位置表現を状態表現とするのではなく，センサ空間から自律的に構成したものを用いる．これにより位置の計算のための観測とそれ以外の意思決定に必要な観測を区別することなく扱うことができる．また，移動ロボットの研究で多く用いられているデッドレコニングのモデルは，一般に場所に依存せず，一定の行動には一定の移動量を対応させるものである．しかし歩行による移動の場合，移動距離がすべりなどのため歩行面へ依存して大きく変化し，移動量のモデルは場所に応じたものが必要である．そのため，あらかじめ移動量のモデルを用意することは困難であり，移動量のモデルが不要であるか，自律的に獲得されることが望まれる．

観測の基準としては，様々な事象を統一的に扱うことができることから，エントロピーを用いる．しかし，従来提案されている状態量推定のための観測戦略ではなく，行動決定のための観測戦略を提案することを目的とする．すなわち，行動決定が可能であれば，自己位置や状態量に曖昧さを状況に応じて許す観測戦略を目指す．

歩行ロボットの移動中の観測へ対応するために，行動決定の曖昧さの指標として，観測により将来得られると期待される情報量を曖昧さの指標として用い，行動確率あるいは情報量期待値が閾値を満足しない場合には，移動せずに観測する戦略をにより，移動中の観測における行動決定の曖昧さの問題に対処する．また，歩行による揺れや，視点移動の影響を軽減する画像補正メカニズムを提案し，センサの外乱，観測間での移動の問題に対処する．画像補正のための補正値の計算には画像列のみを用い，他のセンサは用いない．

補正値を用いて，視覚センサによる観測から，静止時にある観測が得られる確率を計算することで，高速な制御や特別な機構なしに，観測時刻と観測場所の異なる複数観測の統合と揺れの補正を実現することを目指す．

第3章　情報量による観測対象選択

3.1　緒言

本章では，視覚センサを持つ移動ロボットの行動決定のための効率のよい観測戦略を実現する情報量による観測対象選択手法を提案する．観測戦略の効率は行動決定までの観測時間の短さであるとし，明示的な自己位置同定は行わない．観測の行動決定への寄与を情報量により比較し，寄与の大きい観測から実行することにより，効率的な観測戦略を実現する．

以下では，まずランドマークの観測基準としての情報量を用いることにより，より早く行動決定の曖昧さを減少することができることを示す．次にトレーニングデータを圧縮して記憶するために，情報量を基準とした分類木の手法を用いて行動決定木を生成すれば，木の上でのランドマークの順序と情報量を基準にした観測順序が一致することを示す．そして観測戦略に必要な確率計算を示し，行動決定法を述べる．

観測戦略は次のようになる．各行動をとるべき確率を行動確率と呼び，いずれかの行動確率が十分に高くなるまで，情報量が大きいランドマークから観測する．さらに，各ランドマークがある方位に観測される確率を観測確率と呼び，一時刻前の観測確率と，行動から計算する．計算には行動決定木の生成と同じ手法を用いて生成した予測木を用いる．

3.2　情報量による観測対象選択

3.2.1　環境設定

ロボットや環境，与えられるデータ等に関して以下を仮定する．

1) ロボットの視野角が限られており，行動決定に必要な情報が瞬時には得られることが少ない．

2) ランドマークが配置されており，カメラを振り視野を拡大することにより，行動決定に十分な情報が得られる．

3) 行動決定に必要な情報は一定ではなく，状況に応じて変化する．

4) 決定木作成のため，行動，視覚情報は離散化されており，情報量の計算と決定木の生成に十分なデータが用意されている．

本研究の実験では，4) のデータを用意するため教示を用いた．以下，このデータをトレーニングデータと呼ぶ．また説明を簡単にするため，手法の説明においてセンサ情報としてランドマークの方位のみを扱うが，それ以外のセンサ情報も同様に扱うことができる．

3.2.2 観測基準としての情報量

行動の種類を r，ランドマークの方位の分割数 (観測されない場合を含む) を q，ランドマークの種類の数を m，トレーニングデータの数を n とする．各トレーニングデータは，カメラを振った際に観測される各ランドマークの方位と，そのときにとるべき行動を含む．まず各行動 $k = 1, ..., r$ のトレーニングデータ中での生起確率 p_k を求める．行動 k をとった回数を n_k とすると，

$$p_k = \frac{n_k}{n} \tag{3.1}$$

となる．このときの各行動の生起確率の曖昧さを表す量であるエントロピー H_0 は，

$$H_0 = -\sum_k p_k \log_2 p_k \tag{3.2}$$

である．次に，ランドマークの方位がわかった場合の事後生起確率を求める．ランドマーク i が 方位 j に観測されたときに行動 k をとった回数を n_{ijk} とすると，

$$p_{ijk} = \frac{n_{ijk}}{\sum_k n_{ijk}} \tag{3.3}$$

となる．このときの行動に関するエントロピー h_{ij} は，

$$h_{ij} = \sum_k (p_{ijk} \log_2 p_{ijk}) \tag{3.4}$$

である．ランドマーク i が 方位 j に観測される統計的確率 p^l_{ij} は n_{ijk} を使って，

$$p^l_{ij} = \frac{\sum_k n_{ijk}}{\sum_j \sum_k n_{ijk}} \tag{3.5}$$

と表せる．したがって，ランドマーク i の方位を知ることにより期待されるエントロピー H_i と情報量 I_i を計算すると，

$$H_i = -\sum_j p^l_{ij} h_{ij} \tag{3.6}$$

$$I_i = H_0 - H_i \tag{3.7}$$

となる．すなわち，期待される情報量が大きいランドマーク $(\arg\max_i I_i)$ から観測することで，より早くとるべき行動の曖昧さを減少させ行動を決定することができる．

Table 3.1 Example training data of ID3.

no.	attribute A	attribute B	attribute C	class
1	a	b	a	x
2	a	a	b	x
3	b	a	a	y
4	b	a	b	z

3.2.3 決定木，予測木の生成

与えられたデータをいくつかのクラスに分類するためのコンパクトな分類木を生成する方法として ID3 [68] がある．分類木を生成するためには，トレーニングデータセットが必要であり，各データは分類先のクラスと分類のための属性の離散値からなる．属性 i について知ったときのクラスに関する情報量 I_i をすべての属性について計算する．最も情報量の大きい属性について，その属性値によってトレーニングデータセットを分割する．分類木としては，この属性をノードとし，属性値によって枝分かれさせる．新しく分割されたそれぞれのデータセットについて，情報量を計算し，全ての属性について情報量が 0 になるまで，データセットの分割を繰り返すことにより分類木を生成する．その結果，分割後のデータセットに一つのクラスのデータのみが含まれ，その枝の先には，そのクラスを示す葉をおく．

例えば，トレーニングデータセットが Table3.1 の場合には，クラスに関する情報量を計算すると属性 A は 0.5, 属性 B は 1.2, 属性 C は 1.0 となる．そこで属性 A の値 a, b によってトレーニングセットを No. 1,2 と No. 3, 4 に分割する．No. 1,2 のデータセットはクラス x のみであるので，情報量は 0 である．No. 3,4 のデータセットにおいて，クラスに関する情報量を計算すると属性 B は 1, 属性 C は 0 となるので，属性 C によりデータセットを分割し，木の生成を終える．出来上がった木は Fig.3.1 となる．

ここで，分類先のクラスをとるべき行動に，属性を各ランドマークの方位とすることにより，行動決定木を生成する．このように生成された決定木では，木の根から葉に向けてノードは，行動決定に関する情報量の大きいランドマークから順に並んでいる．したがって，観測を情報量の順に行うことは，木の根に近いランドマークから観測することに等しくなる．ランドマーク i の予測木は，一時刻後のランドマーク i の方位を分類先クラスとし，現時刻での各ランドマークの方位と，行動を属性とすることにより生成する．

実際のトレーニングデータには，センサ値を離散化しているため，同じ観測で異なる行動をとったデータが含まれる場合がある．その場合に関しては，それぞれのトレーニングデータ中でそれぞれの行動をとった確率を計算し別の葉として生成する．ランドマーク予測木の生成に関しても同様である．

Table 3.2 のトレーニングデータが得られた場合の計算は次のようになる．まず，$p_x = 0.2$,

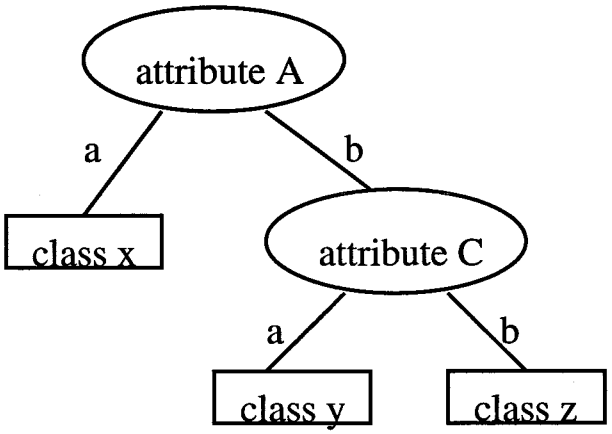


Fig.3.1 Example classificatoin tree by ID3.

Table 3.2 An example of teaching data.

landmark A	landmark B	landmark C	action
a	a	a	x
b	a	a	y
c	b	a	y
a	c	a	z
a	b	b	z

$p_y = 0.4, p_z = 0.4$ より，$H_0 = 1.52$ となる．p_{ijk} を計算し，情報量を計算すると $I_A = 0.97$, $I_B = 0.72$, $I_C = 0.32$ となるので，決定木でのランドマークの順位を A, B, C とし，木を生成すると Fig.3.2 となる．

3.2.4 確率の計算

時刻 t での各ランドマーク i がどの方位 j に観測されるかの確率 (観測確率) を $p_{ij}^L(t)$ $(i = 1, ..., m, j = 1, ..., q)$ とし，時刻 t で行動 k をとった確率を $p_k^a(t)$ $(k = 1, ..., r)$ とする．過去の経験から時刻 t でとるべき行動が k である確率 (行動確率) を $\hat{p}_k^a(t)$ $(k = 1, ..., r)$ とする．

確率の計算は次のように行う．現在画像上で観測されているランドマーク i については，その方位 J の確率を $p_{iJ}^L(t) = 1$ とし，それ以外を $p_{ij}^L(t) = 0 (j \neq J)$ とする．1 時刻前の行動については，実際にとった行動 K を $p_K^a = 1$ とし，それ以外を $p_k^a = 0 (k \neq K)$ とする．観測されていないランドマークについては，1 時刻前のそれぞれの確率 $p_{ij}^L(t-1)$ から予測木を使って計算する．カメラを振ることによりランドマーク探索した場合にも観

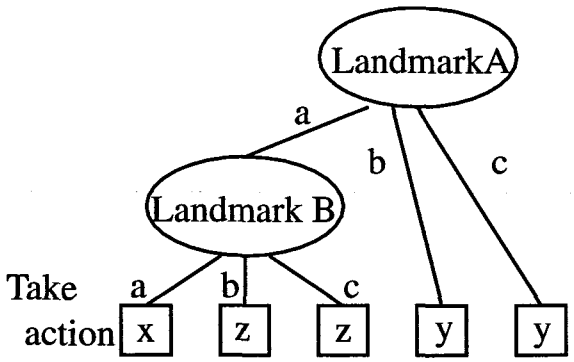


Fig.3.2 An example of an action decision tree.

測されない場合にのみ，観測されていない確率を 1 とし，残りの方位に観測される確率を 0 とする．

予測木からの観測確率計算は次のように行う．ランドマーク i の木の根から葉までたどると，時刻 $(t-1)$ の各ランドマークの方位と行動の論理積を満たした場合の，ランドマーク i の時刻 t での方位が記述されている．そこですべての葉について，論理積をその方位であった (行動をとった) 確率の積に置き換え，その葉に到達する確率の計算を行う．複数の葉に同じ方位が現れるので，それらの和を時刻 t での，その方位となる確率と見なす．例えば，ランドマーク A の予測木が Fig.3.3 であれば，

$$\begin{aligned} p^L_{Aa}(t) &= p^a_x(t-1)p^L_{Ab}(t-1)p^L_{Bb}(t-1)\times 0.3 \\ p^L_{Ab}(t) &= p^a_x(t-1)p^L_{Aa}(t-1)p^L_{Ba}(t-1)p^L_{Cb}(t-1) \\ p^L_{Ac}(t) &= p^a_x(t-1)p^L_{Ab}(t-1)p^L_{Ba}(t-1)p^L_{Cc}(t-1) \\ &\quad + p^a_x(t-1)p^L_{Ab}(t-1)p^L_{Bb}(t-1)\times 0.7 \end{aligned} \tag{3.8}$$

となる．

これらから得られた時刻 t でのランドマークの方位の観測確率 $p^L_{ij}(t)$ を用いて，行動決定木を同様にたどり，行動確率 $\hat{p}^a_k(t)$ を計算する．

3.2.5 行動決定

行動確率の計算後，行動を決定する．ある行動の確率が，特に高ければその行動をとればよい．そうでなければ，特定の行動確率が十分に高くなるまで，行動決定木の上から順にランドマークの観測確率を調べ，方位の曖昧なランドマークについて再観測を繰り返す．すなわち情報量基準で順に再観測の必要なランドマークを調べる．また再観測の際，ランドマークの観測確率の高い方位を優先的に調べることで，再観測時間を軽減できると期待される．

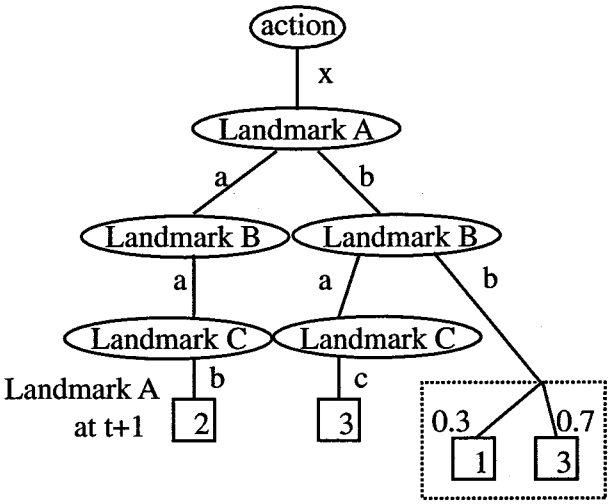


Fig.3.3 An example of a prediction tree for landmark A.

3.3 実験

3.3.1 実験環境

ロボットとしては，RoboCup 99 SONY 脚式ロボットリーグ [69] のロボット (Fig. 3.4) を用いた．カメラはロボットの頭部の鼻に当る位置に設置されており，画角は横53度，縦41度，画素数はそれぞれ88, 59である．脚は各3自由度，カメラは3自由度(パン，チルト，視線周り)あるが，ランドマークを観測する際には脚の角度と，カメラのチルト，視線周りの角度を固定し，パン軸のみを利用した．パン軸はロボット正面に対して，-90度から90度が可動範囲である．

環境をFig.3.5に示す．ランドマークは8個あり，ボールが一つある．それぞれ，敵ゴール (TG), 自陣ゴール (OG), 北西ポール (NW), 北東ポール (NE), 中央西ポール (CW), 中央東ポール (CE), 南西ポール (SW), 南東ポール (SE) とする．すべてのランドマークとボールは色により識別される．ロボットがボールをTGに入れることをタスクとする．これを実現するためには場所に応じたボールへの回り込み，ボールの探索などが必要となる．

各ランドマークの方位は，ロボットに対して前方を0度として，$(,-65°)$, $[-65°,-40°)$, $[-40°,-15°)$, $[-15°,15°)$, $[15°,40°)$, $[40°,65°)$, $[65°,)$ の7方位に分割し，観測されない場合(oで表す)を含めて方位は8通りとした (Fig. 3.6 参照)．ボールの方向は，$(,-45°)$, $[-45°,-12°)$, $[-12°,12°)$, $[12°,45°)$, $[45°,)$ の5方位に分割し，さらにロボットから遠い近いの2通り(水平に対して下向30度で分割)に分け，観測されない場合(oで表す)を含めて方向は11通りとした．ボールは，ロボットの行動と一時刻前のボールの方向のみに依存する特殊なランドマークとして扱い予測木を生成した．

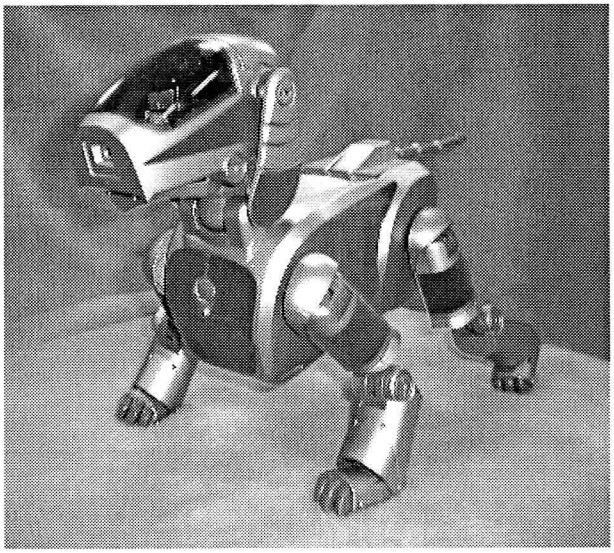
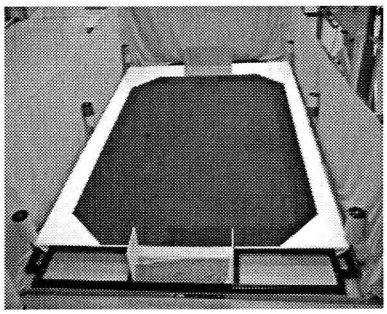
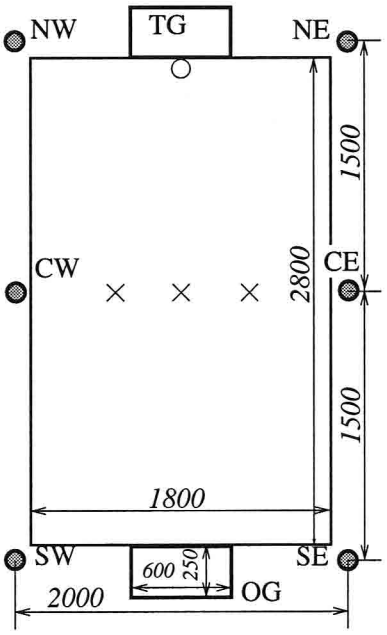


Fig.3.4 The SONY legged robot for RoboCup 99 SONY legged robot league.

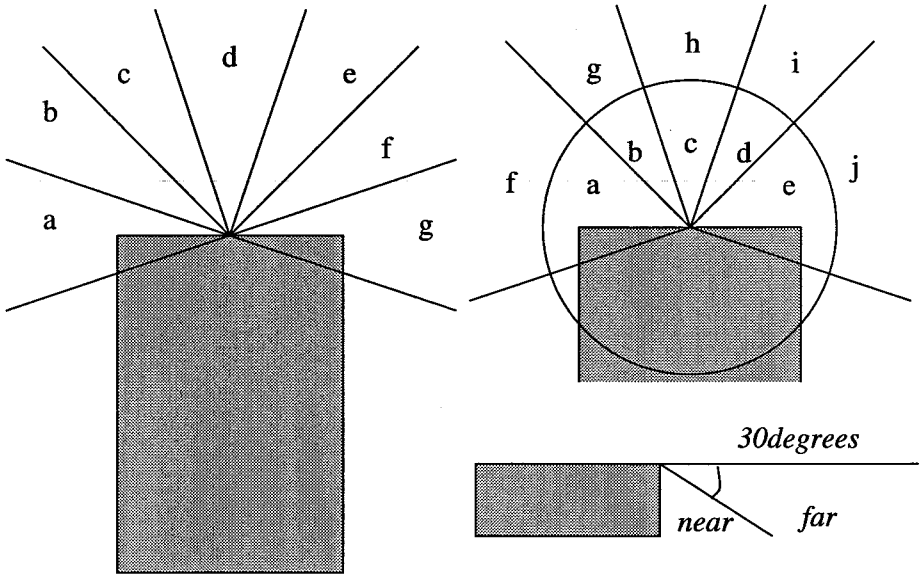


(a) Photo of the field.



(b) Size of the field

Fig.3.5 Experimental field (same as the one for RoboCup SONY legged robot league). Cross and circle marks are for the first experiment.



(a) for landmarks (b) for the ball

Fig.3.6 Quantization for landmarks and the ball.

行動決定木や各予測木をたどって確率を計算する際，各木を生成する時に含まれなかった方位により，確率の合計 $\sum_{i=1}^{N} p_i$ が1にならない場合がある．ここでは，合計が1となるよう $(1-\sum_{i=1}^{N} p_i)/N$ を $p_i(i=1,...,N)$ に加えた．N は行動の種類数あるいは方位の分割数である．

生成された木を使って行動を決定する (以下教示再生と呼ぶ) 際には，行動確率の最大値が 0.6 以上であればその行動をとり，そうでなければカメラを振って観測されていないランドマークとボールを再探索することとした．教示中，教示再生時ともにランドマーク再探索時以外は，ボールを追跡視あるいは探索するようにした．

3.3.2 実験 1

まずFig.3.5のフィールドにおいて，ゴール前 (図の○印) にボールをおき，フィールドの中央3点 (フィールド中央の3つの×印) から開始してボールをゴールにいれるタスクを行った．行動は我々の開発した歩行プログラムを用い，前進，左右大回りの3つとした．行動の継続時間は，4.8秒とした．これは8歩行周期で，直進の場合約 0.45[m] の移動となり，ほぼ一度の行動で観測が変化するからである．教示は中央3点から各5回ずつ行い，80のデータを得た．このデータから生成した決定木の一部をFig.3.7に示す．ボールが左前方 (b の方向) に観測された場合には，回り込みのため敵ゴールの方位 (c,d,f) に応じて

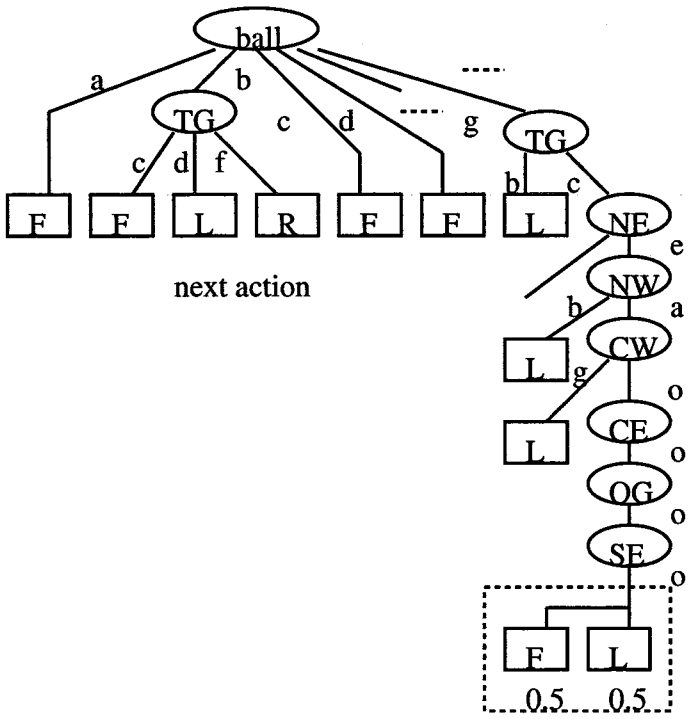


Fig.3.7 A part of the action decision tree (Experiment 1). F, L, and R mean forward, left forward, and right forward respectively.

とるべき行動が変化することが分かる．決定木と予測木の大きさと情報量による順序をTable 3.3 から 3.5 に示す．行動決定木の大きさは，葉の数が 43，最小深さは 1，平均深さは 4.91，最大深さは 8 となった．Table 3.3 は予測木の大きさと深さを示し，表中 # of leaves は葉の数を表し，min dep., mean dep., max dep. は，それぞれ木の最小深さ，平均深さ，最大深さを表す．進行方向と逆にあり，小さく観測されることの少ない SW, SE などのランドマークの予測木は小さくなっている．Table 3.4 (Table 3.5) は行動決定木(各予測木)における，ランドマークなどの情報量による順位を表す．1 が最も木の根に近く情報量が大きい，8 が最も根から遠く情報量が小さい．Table 3.4, 3.5 で ball は一時刻前のボールの方向を，act は一時刻前の行動を，TG, OG, NW, NE, CW, CE, SW, SE はそれぞれのランドマークの一時刻前の方位である．行動決定木の最上位にボールが，予測木の比較的上位に行動が現われている．これらの木を使った教示再生中の動きを Fig.3.8 に示す．

次に教示再生を行った場合の予測と実際にとった行動例を示す．まずフィールド中心から開始した場合に，ロボットは，1) 前進→ 2) 前進→ 3) 前進→ 4) 前進という行動をとった．ボールと TG は常に観測されており，各行動決定時の観測確率と行動確率は Fig.3.9 のようになった．各マス目の濃度は確率を表し，白が 0，黒が 1 である．ball はボール

Table 3.3 Depth of the prediction trees(Experiment 1).

tree for	# of leaves	min dep.	mean dep.	max dep.
ball	52	2	2	2
OG	13	1	4.23	8
TG	44	1	5.39	8
SE	6	1	2	3
SW	1	0	0	0
CE	28	2	4.69	8
CW	11	1	3.91	8
NE	51	1	5.96	8
NW	54	2	5.91	8

Table 3.4 The order of information for the action decision tree (Experiment 1).

1	2	3	4	5	6	7	8
ball	TG	NE	NW	CW	CE	OG	SE

Table 3.5 The order of information for prediction trees (Experiment 1).

tree for	1	2	3	4	5	6	7	8
ball	ball	act						
OG	act	NE	TG	NW	CW	CE	OG	SE
TG	TG	act	NE	NW	CE	OG	CW	SE
SE	act	CE	NE	OG	NW	TG	CW	SE
SW	-							
CE	act	NE	TG	CE	NW	CW	OG	SE
CW	TG	act	NE	NW	CE	CW	OG	SE
NE	NE	act	NW	TG	CE	CW	OG	SE
NW	act	NE	TG	NW	CE	OG	SE	CW

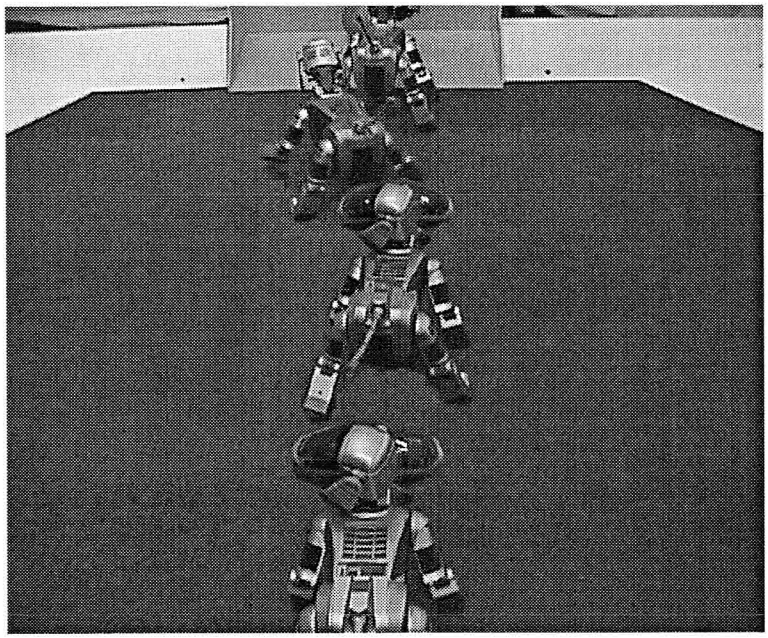


Fig.3.8 Robot movements with the action and decision trees.

の方向の，OG, TG, SE, SW, CE, CW, NE, NW はそれぞれのランドマークの方位の予測木あるいは実際の観測による観測確率であり，not seen はカメラを振っても観測されない場合の確率を，それ以外はボールは下段の左から a,b,...,e の上段の左から f,g,...,j の確率を，ランドマークは左から順に方位 a,b,...,g の確率を表している．action は行動決定木による行動確率であり，F, L, R はそれぞれ．前進，左大回り，右大回りをとるべき確率を表している．行動開始時はカメラを振って観測を行っている．Fig.3.9 から，時刻 3, 4 ではいくつかのランドマークの観測確率が平均して低いものになっているが，カメラを振ることなく行動決定を行えている．

同じフィールド中心から開始した場合でも，別の行動をとることもあった．これは，初期位置のロボットからのランドマークの観測や歩行の結果が必ずしも同一ではないからである．この例では，初期のランドマークの観測された方位は一致しているが，歩行の結果が一致しなかった．ロボットは，1) 前進→ 2) 前進→ 3) ランドマーク確認→ 4) 前進→ 5) ランドマーク確認→ 6) 左回転→ 7) 前進→ 8) 前進という行動をとった．ボールと TG は常に観測されており，観測確率と行動確率は Fig.3.10 のようになった．時刻 3, 5 では行動決定ができず，首を振った観測を行う必要が出ている．一方，時刻 7, 8 では観測確率が曖昧なランドマークがあっても行動決定が行われている．

次にフィールドの右側から開始した場合の行動例を示す．ロボットは，1) 左回転→ 2) ランドマーク確認→ 3) 前進→ 4) 前進→ 5) 左回転という行動をとった．ボールと TG は常に観測されており，観測確率と行動確率は Fig.3.11 のようになった．時刻 2 では行動確

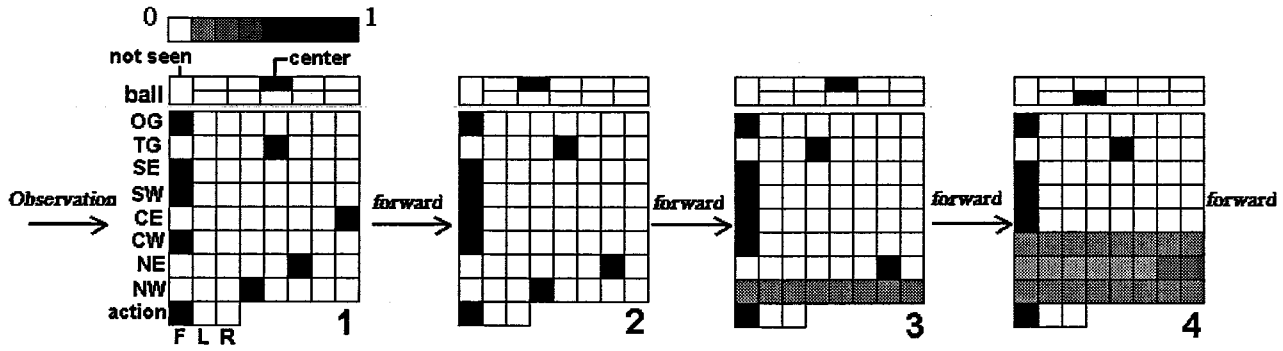


Fig.3.9 Probability distribution in Experiment 1-1 (The gray level of each box indicates the probability, black 1 and white 0) .

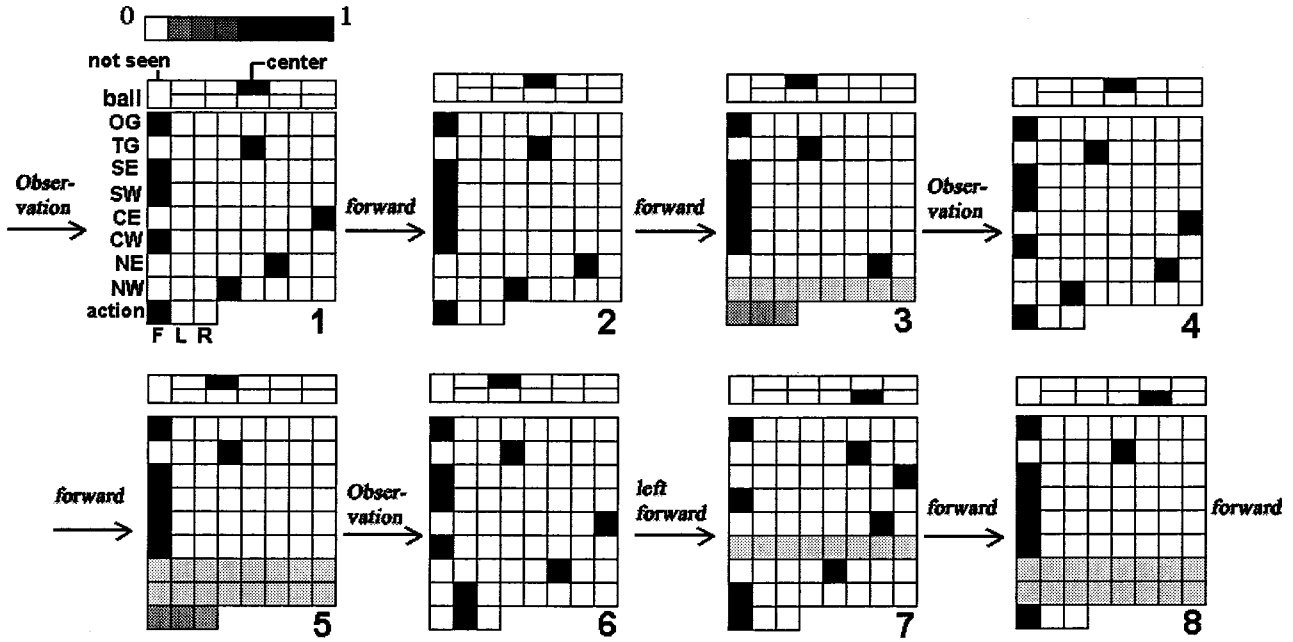


Fig.3.10 Probability distribution in Experiment 1-2.

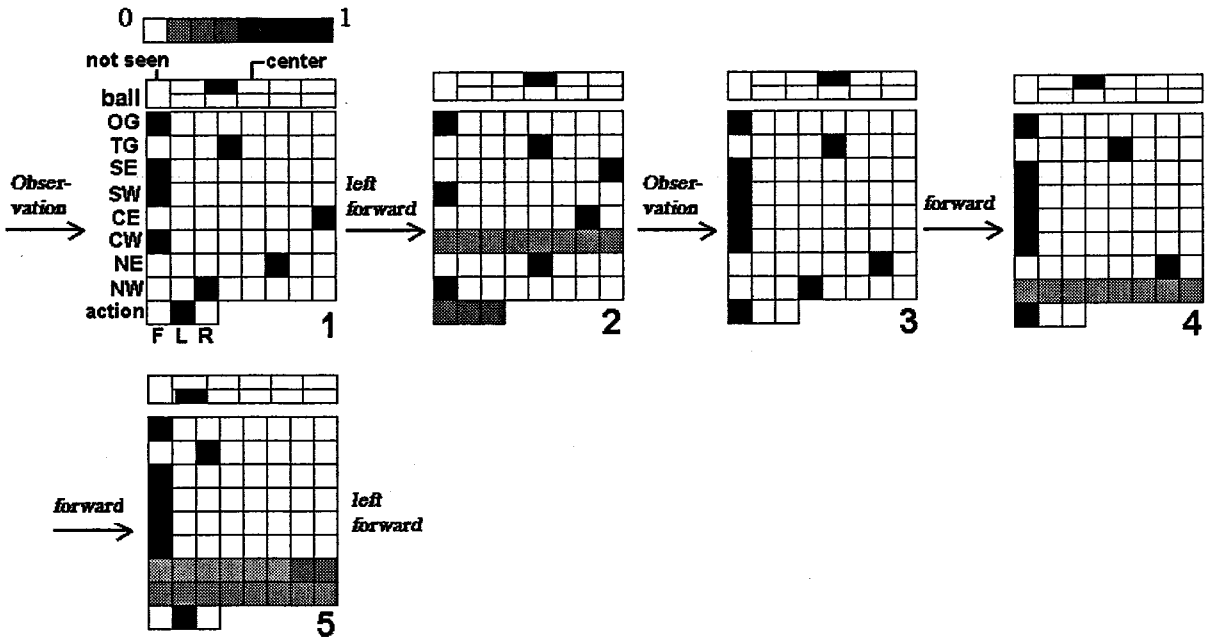


Fig.3.11 Probability distribution in Experiment 1-3.

Table 3.6 The number of needed observation (Experiment 1).

begin from	# of trials	# of total steps	# of re-observations	rate of re-observation
center	12	35	18	.51
left	12	31	15	.48
right	16	64	38	.59

率が曖昧なため行動決定ができず，時刻 4, 5 では観測確率が曖昧なランドマークがあっても行動決定が行われている．

教示再生中の再観測回数を Table 3.6 に示す．左から順に試行開始場所，試行回数，合計行動回数，合計再観測回数，再観測率である．再観測回数が半分程度に減少していることが分かる．

3.3.3 実験 2

同じフィールド上で RoboCup 99 の試合を考慮した教示を行った．教示の負担を減らすため，行動は，前進，左右大回り，左右小回り，ボール追跡の 6 つとした．行動の継続

Table 3.7 Depth and size of the action decision tree (Experiment 2).

# of leaves	min dep.	mean dep.	max dep.
586	2	5.89	9

Table 3.8 Depth and size of the prediction trees (Experiment 2).

	# of leaves	min dep.	mean dep.	max dep.
ball	403	2	2	2
OG	958	2	7.58	9
TG	1050	2	7.67	9
SE	845	2	7.35	9
SW	901	2	7.41	9
CE	901	2	7.13	9
CW	873	2	7.37	9
NE	1031	2	7.60	9
NW	980	2	7.55	9

時間は，前の実験と同じ4.8秒とした．この教示により，1364のデータを得た．このうち不適切な教示を除き，856を行動決定木の生成に，1364すべてを予測木の生成に用いた．各木の大きさ，情報量による順序をTable 3.7から3.10に示す．ランドマーク毎の予測木の平均深さの差が少なくなっていること，行動決定木の上位二つは変らないこと，予測木の最上位がほぼ予測対象のランドマークとなっていることなどが分かる．

このデータを実際に，RoboCup 99で用いたところ，ロボットは教示者の期待した行動を行った．しかしながらランドマークの確認が頻繁に行われた．これはランドマークの予測木がうまく働かなかったためと思われる．

Table 3.9 The order of information for the action decision tree(Experiment 2).

1	2	3	4	5	6	7	8	9
ball	TG	OG	SW	SE	NW	NE	CE	CW

Table 3.10 The order of information for prediction trees (Experiment 2).

tree for	1	2	3	4	5	6	7	8	9
ball	ball	act							
OG	OG	SE	SW	TG	NW	CW	NE	CE	act
TG	TG	OG	SE	SW	NW	NE	CW	CE	act
SE	SE	OG	TG	SW	CE	NE	NW	CW	act
SW	SW	OG	CW	SE	TG	NW	NE	CE	act
CE	CE	SE	OG	TG	NE	SW	NW	act	CW
CW	CW	SW	OG	TG	NW	SE	NE	CE	act
NE	TG	NE	OG	SE	CE	NW	SW	CW	act
NW	NW	TG	OG	SW	CW	SE	NE	CE	act

3.4 討論

実験 1 の結果 (Fig.3.9-3.11) から，行動確率はほぼ 1 の行動があるか，ほぼ一様な確率となるかのいずれかとなっている．行動確率が一様な場合，観測確率の低いランドマークの確率も平均して低くなっている．このため観測確率の高い方位を優先的に観測することはできていない．また，行動確率が低かったため，カメラを振り観測を行った前後の観測確率を比較すると，確率が平均して低かったランドマークの確率が変化するだけでなく，より情報量の大きいランドマークの確率も変化している．原因としては，予測木の作成に用いたトレーニングデータが少なかった，あるいは予測木の汎化能力が不足している，またランドマークの予測には分類木は向かないといった問題が考えられる．

実験 1, 実験 2 のランドマーク予測木を比較すると，実験 1 では行動が上位に来ており，実験 2 では行動が下位になっている．実験 2 の結果のように場所に依存して観測される方位の異なるランドマークの予測は，ランドマークにより大まかな位置が判明してから行った方がよい．実験 1 でこれが見られないのは，トレーニングデータが少ないためであると考えられる．

ここでは，決定木の圧縮はさほど行っていない．圧縮を行うとトレーニングデータにない状況への対応が期待される一方，ランドマークの再確認を行うべき状況において確認を行わないことが増加すると考えられる．また各ランドマーク，ボールの方向の離散化方法をここではあらかじめ決定している．しかし今回利用した離散化法が最適であるとはいえない．連続な属性値を情報量基準により閾値によって2分していくことにより扱う C4.5[70] のような分類木生成法を行動決定木の生成に用いることで決定木の生成と自律的な離散化を同時に行える可能性がある．

この手法は，to look or to move また，what to look の解にはなると思われるが，行動

の時間的区切り方(ここでは4.8秒で固定), when to look の問題は残っている. さらに, 一度見渡せば場所に関して十分な情報を得られると仮定しているが, 仮定の成立しない場合への対処が必要である. また, 行動中の注視対象をボールに限ったが, これも情報量基準で選択することが望ましい. ここでの実験では行動確率から行動を決定する際の閾値は実験的に決定したが, 行動確率からの行動決定法とともに閾値の適切な決定法は今後の課題である. データ収集は教示により行ったが, 教示者の負担も大きく, 負担の軽いデータ収集法が望まれる.

3.5 結言

本研究では, 移動体の意思決定のための効率的な観測のための情報量を利用した観測戦略を提案した. さらに行動決定木を情報量をもとに作成することにより, トレーニングデータの圧縮と観測順序の記述が統一的に扱えることを示した. そして実機実験により本手法の有効性を検証した. 提案した手法は, 視覚センサによるランドマークの観測のみでなく, 他の能動センサによる観測も統一して扱うことが可能である.

第4章　情報量による観測対象と注視窓の選択

4.1　緒言

前章で，自己位置の同定や状態推定ではなく行動決定のための，効率的な観測を行う行動決定法を提案した．この手法では情報量による観測対象選択を行っているが，観測方向についてはセンサ値の予測に依存していた．またセンサ値が前もって離散化されており，観測にかかる時間は一定であると仮定していた．しかし，さらに効率的な観測を行うには，観測対象と観測方向を考慮した注視制御のための自律的センサ空間分割が必要となる．

そこで本章では，注視制御のためのセンサ空間の分割と行動決定木の生成法を提案する．自己位置の同定のための観測ではなく，観測にかかる時間を考慮した行動決定のための効率的な観測を実現する．また履歴の考慮のため行動決定木のを再生成を提案する．

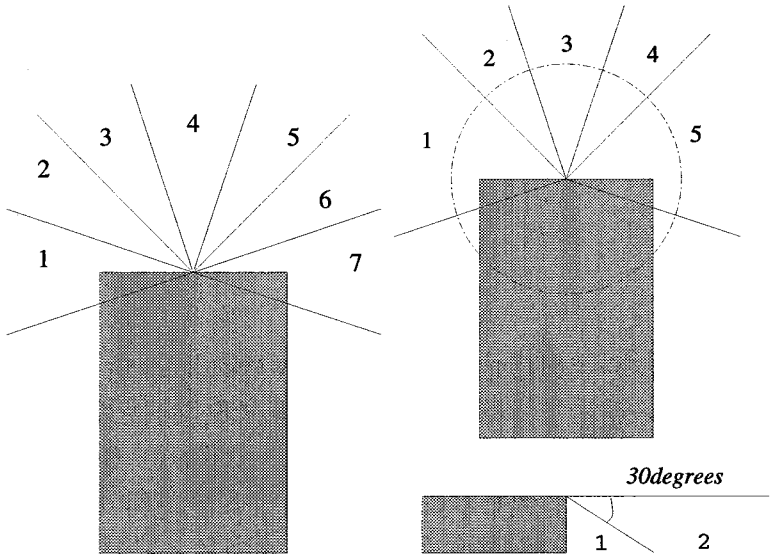
4.2　自律的センサ空間分割の重要性

前章で提案した手法ではセンサ値は離散値をとるものとして扱っていた．センサ値は連続した値であることが多いが，離散値とするために予めセンサ空間を分割していた．これには，離散値として扱うことにより数値の連続性 (大小関係) が損なわれる，必要なセンサの解像度を予め知る必要がある，といった問題がある．情報量による決定木の生成法である C4.5 [70] では，属性が連続値である場合には，センサ値と閾値の大小関係でデータセットを 2 分割した際に，情報量が最も大きくなる閾値を求め，閾値でデータセットを 2 分割することで，トレーニングデータの分割を行う．これにより，決定木の生成にセンサ値の連続性を利用できると期待される．

前章の実験 2，RoboCup 99 の試合を考慮した教示のデータを用いて，C4.5 のトレーニングデータ分割により行動決定木の変化を調べた．教示データは，ランドマーク，ボール共に観測されない場合には 0 を，観測される場合には方位に応じて **Fig.4.1** に示す値をとるものとした．ボールについては左右と遠近の二つのセンサがあるものとした．生成された行動決定木と予測木の葉の数を **Table 4.2** に示す．教示データ数は 856 である．

ボール予測木をのぞき，行動決定木，ランドマーク予測木は，C4.5 を用いることで小さくなっていることが分かる．これはセンサ値の連続性の利用によると考えられる．ボー

ル予測木で木が小さくならないのは，ボールはロボットの行動により移動する，利用しているセンサがボールの左右と遠近の2つと行動のみであるためと考えられる．



(a) for landmarks (b) for the ball

Fig.4.1 Quantization for landmarks and the ball.

	action	OG	TG	SE	SW	CE	CW	NE	NW	ball
previous	557	510	610	472	489	507	482	585	585	258
C4.5	419	353	466	308	328	341	311	472	423	256

Table 4.1 Comparison of the number of leaves between the previous method and the C4.5

4.2.1 視覚センサの特性と注視窓

視野角の限られた視覚センサの場合では，視野角外の情報は当然ながら得ることができない．したがって前章のように，観測をあるランドマークの方位を知ることとし，行動に関する情報量を計算して生成した行動決定木の場合には，直接観測方向を決定することはできず，予測木などを用意する必要がある．一方，ある方向を観測した場合の情報量を計算し，観測方向の決定に利用することが出来れば都合がよい．たとえば，ある方向に特定

のランドマークが観測されるか否かを知ることによる情報量を比較し，観測方向と観測対象を決定することが考えられる．

またランドマークの図心座標など値が連続値である場合も扱いたい．しかし，ランドマークの図心座標と閾値の大小関係を知るためには，ランドマークが既に観測されている場合を除き，視覚センサの方向を何度か変更しランドマークを観測する必要がある．したがって，C4.5 で用いられている閾値による分割方法を用いると観測方向を一意に決定できない．それに対し，ランドマークがある範囲に観測されるか否かを知るためには，その範囲を含む方向の一度の観測で済む．そこで，ある範囲 (以下，注視窓と呼ぶ) 内に特定のランドマークが観測されるか否かを知ることによる情報量を比較し，観測方向 (注視窓を含む方向) と観測対象を決定することにより，自律的センサ空間の分割と行動決定木の生成を同時に手法を提案する．この行動決定木により，観測対象と観測方向が決定される．

4.3 情報量に基づく効率的観測と行動決定

4.3.1 環境設定

ロボットや環境，与えられるデータ等に関して以下のように設定する．

1. ロボットは視覚センサを持つが視野角が限られており，受動的な観測のみでは行動決定に必要な情報が得られるとは限らない．

2. 環境中にランドマークが配置されており，視覚センサの観測方向を変えることにより，視野を拡大し行動決定に十分な情報が得られる．

3. 教示などにより，視野を拡大した際に得られる視覚センサの値と，その際にとるべき行動のシーケンスが与えられる．

また行動決定に必要な観測情報は一定ではなく，状況に応じて変化する．

4.3.2 観測による情報量

行動の種類を r，トレーニングデータの数を n とする．トレーニングデータ中の行動 j をとった回数を n_j とすると，各行動 $j = 1, ..., r$ の生起確率 p_j は，$p_j = n_j/n$ である．このとき行動に関するエントロピー H_0 は，次のようになる．

$$H_0 = -\sum_{j=1}^{r} p_j \log_2 p_j \tag{4.1}$$

ランドマーク i が注視窓 $[\theta_{Lk}, \theta_{Uk})$ の範囲に観測されるか否かが分かったときの事後生起確率を求める．ランドマーク i が $[\theta_{Lk}, \theta_{Uk})$ に観測された回数を n_{ik}^I，観測されたとき

に行動 j をとった回数を n^I_{ikj}, $n^I_{ik} = \sum_{j=1}^{r} n^I_{ikj}$ とすると，注視窓内に観測される場合の事後生起確率は，

$$p^I_{ikj} = \frac{n^I_{ikj}}{\sum_j^r n^I_{ikj}} = \frac{n^I_{ikj}}{n^I_{ik}} \tag{4.2}$$

となる．同様に，観測されなかった場合の行動 j をとった回数を n^O_{ikj}, $n^O_{ik} = \sum_{j=1}^{r} n^O_{ikj}$ とし，観測されない場合の事後生起確率 p^O_{ikj} を求める．$n_{ik} = n^I_{ik} + n^O_{ik}$ として，このときのエントロピーを計算すると，

$$H_{ik} = -\sum_{x=\{I,O\}} \frac{n^x_{ik}}{n_{ik}} \sum_{j=1}^{r} (p^x_{ikj} \log_2 p^x_{ikj}) \tag{4.3}$$

となり，それぞれの観測による情報量は，$I_{ik} = H_0 - H_{ik}$ である．情報量が大きいランドマーク i, 注視窓 k ほど，行動に関する曖昧さが減少する．観測範囲の上下限 θ_{Lk}, θ_{Uk} は，各ランドマーク i についてトレーニングデータ中に含まれる隣接する観測された方位の中間点などを候補とする．

4.3.3 観測時間の考慮

観測に要する時間が観測対象 (ランドマークとその観測される範囲) によらず一定の場合には，情報量の最も大きい観測対象により，行動決定木を生成する．この決定木はコンパクトであり，木のノードの観測を繰り返すことで最短観測時間で行動を決定できる．しかし，観測時間が観測対象により異なる場合には，最短観測時間となるとは限らない．また観測対象決定時に，得られる情報量と観測時間のトレードオフを計算するのは，観測時の計算コストが高く，決定木を使うメリットが減少する．

そこで決定木の生成の際に，観測時間を考慮した指標を用いる．現在の観測対象の次にその観測対象を観測する場合にかかる時間を T とし，単位時間当りに得られる情報量 (情報量速度)i_{ik} を次のように求める．

$$i_{ik} = \frac{I_{ik}}{T + T_C} \tag{4.4}$$

ここで T_C は0で割らないための正の定数である．現在，観測していない場合には，行動決定時の平均的なカメラの方向からの観測時間を T として用いる．すでに視覚センサが観測した方向である場合には $T = 0$ とする．

4.3.4 行動決定木の生成

行動決定木を次のように生成する．

1. 情報量速度 i_{ik} を全てのランドマーク i, 注視窓候補 k について計算する．

Table 4.2 Example training data

data number	landmark A	landmark B	action
1	5	5	x
2	25	15	x
3	27	10	y
4	40	30	z

2. i_{ik} を最も大きくする i, k の組により，トレーニングデータを，ランドマーク i が注視窓 $[\theta_{Lk}, \theta_{Uk})$ の範囲に観測された場合と，観測されなかった場合に分ける．

3. 行動が決定できるまで情報量速度の計算とトレーニングデータの分割を繰り返す．

トレーニングデータの分割が木の枝分かれとなる．トレーニングデータに対し最後まで分割を繰り返しても，行動が確定しない場合には，各行動の頻度確率を記しておく．

たとえば，Table 4.2 のトレーニングデータが与えられたとする．表中の数字は各ランドマークが観測された方向である．視野が限られており，視覚センサによる観測可能な方向が $[0,15), [15,30), [30,45)$ の 3 つである．視覚センサは観測開始時に $[15,30)$ を向いており，観測方向を変えるのに 1 時刻必要であるとする $(T_C = 1)$．まず，H_0 を計算すると，$p_x = 2/4, p_y = 1/4, p_z = 1/4$ から，$H_0 = 1.5$ となる．ランドマーク A, B がある範囲に観測されるか否かがわかったときの，情報量 I_{ik} と単位時間当りに得られる情報量 (情報量速度) i_{ik} を計算すると Table 4.3 となる．ここで注視窓の上下限の候補には，各ランドマークの観測された方位の中間点と，観測可能な範囲の上下限を用いている．これは注視窓が各方向を観測した場合の視野を越えないためである．そこで，最も情報量速度が大きいランドマーク A が注視窓 $[26,30)$ に観測されるか否かを確かめる．観測された場合のトレーニングデータはデータ番号 3 のみで，行動は y が決定できる．観測されない場合には，データ番号 1,2,4 が含まれ，行動は決定できない．このトレーニングデータのサブセット (データ番号 1,2,4) で，ランドマーク A, B がある範囲に観測されるか否かがわかったときの情報量速度を計算すると，ランドマーク B が $[0,15)$ に観測されるか否かが分かったときと，ランドマーク A が $[30,40)$ に観測されるか否かが分かったときが最も大きく，共に 0.05 である．左側を優先して観測すると行動決定木は，Fig. 4.2 となる．

4.3.5 行動決定

行動を次のように決定する．

1. 観測予測モデルが使える場合には，モデルを使い各注視窓にランドマークが入る確

Table 4.3 Information gain and information gain per time calculated from example training data (Info., Info./time and Lm indicate information gain, information gain per time and landmark respectively).

observation	info. (I_{ik})	info./time (i_{ik})
$0 \leq (LmA) < 15$	.31	.15
$15 \leq (LmA) < 26$	.31	.31
$15 \leq (LmA) < 30$	.50	.50
$26 \leq (LmA) < 30$	1.4	1.4
$30 \leq (LmA) < 45$	1.4	.70
$0 \leq (LmB) < 7$	.31	.15
$0 \leq (LmB) < 15$	.50	.25
$7 \leq (LmB) < 15$	1.4	.70
$15 \leq (LmB) < 30$	.31	.31
$30 \leq (LmB) < 45$	1.4	.70

率を計算する[1]．モデルを持たない場合には，各注視窓に入る確率 (観測確率) は 0.5 とする．

2. 現在，視覚センサが向いている方向について，各注視窓にランドマークが入っているか，否かがわかるので，観測確率を 1 または 0 とする．

3. 観測確率を用いて行動決定木の各葉への到達確率を計算する．同じ行動を示す複数の葉への到達確率の和を，その行動をとるべき確率とする．

4. 特定の行動が閾値を越えていれば，その行動をとる．そうでなければ，根に近い観測範囲を含むまだ観測していない方向から順に観測し，観測確率を更新する ((2) に戻る)．

更新した確率分布により，特定の行動をとるべき確率が閾値を越えるまで，観測と確率分布の更新を繰り返す．

例えば，Fig. 4.3(a) の行動決定木により行動決定する場合には，次のようになる．行動確率の閾値は 0.8 とする．

1. 観測モデルを持たないので各注視窓にはいる確率は 0.5 であり，葉への到達確率は

[1]例えば，観測予測モデルによりランドマーク i の観測方向が確率分布 $f(\boldsymbol{x})$ に従うと予測されたとする．このとき注視窓 k 内にランドマーク i が観測される確率 $\hat{p}_{ik}$ は，$\hat{p}_{ik} = \int_{W_{ik}} f(\boldsymbol{x})d\boldsymbol{x}$ となる．ここで W_{ik} は注視窓内の座標を表す．ただし，本論文における実験では観測予測モデルを使用せず，観測前の確率は 0.5 とした．

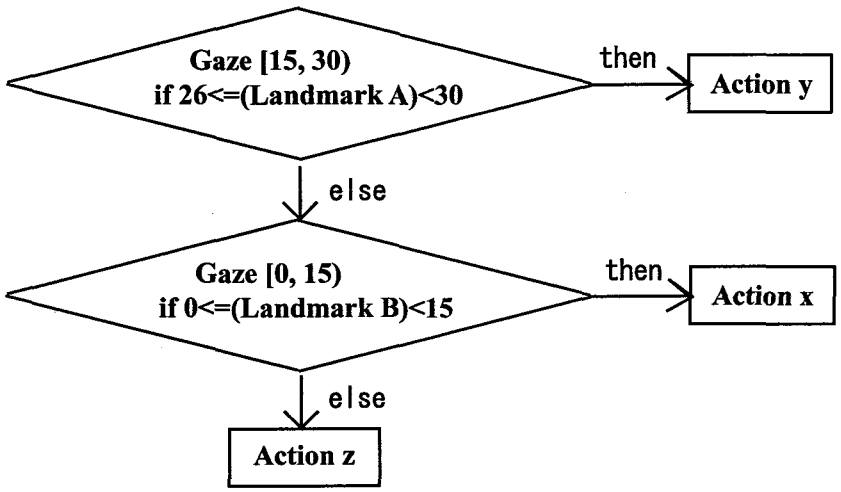


Fig.4.2 The action decision tree constructed from example training data

Fig. 4.3(a) となる．行動確率は，行動 x が 0.5，行動 y が 0.25，行動 z が 0.25 である．

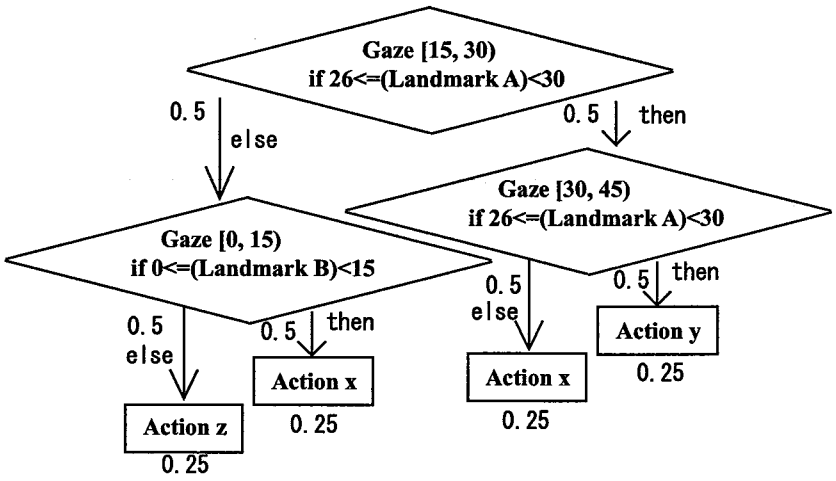
2. 行動確率が閾値を越えていないので，根の注視窓 $[26, 30)$ を含む範囲 $[15, 30)$ を観測し，観測確率を更新する．注視窓内にランドマーク A が観測されたとすると，葉への到達確率は Fig. 4.3(b) となる．行動確率は，行動 x が 0.5，行動 y が 0.5，行動 z が 0.0 である．

3. 行動確率が閾値を越えていないので，次に注視窓 $[35, 40)$ を含む範囲 $[30, 45)$ を観測し，観測確率を更新する．注視窓内にランドマーク B が観測されなかったとすると，葉への到達確率は Fig. 4.3(c) となる．行動確率は，行動 x が 1.0，行動 y が 0.0，行動 z が 0.0 となるので，行動決定を終了し，行動 x をとる．

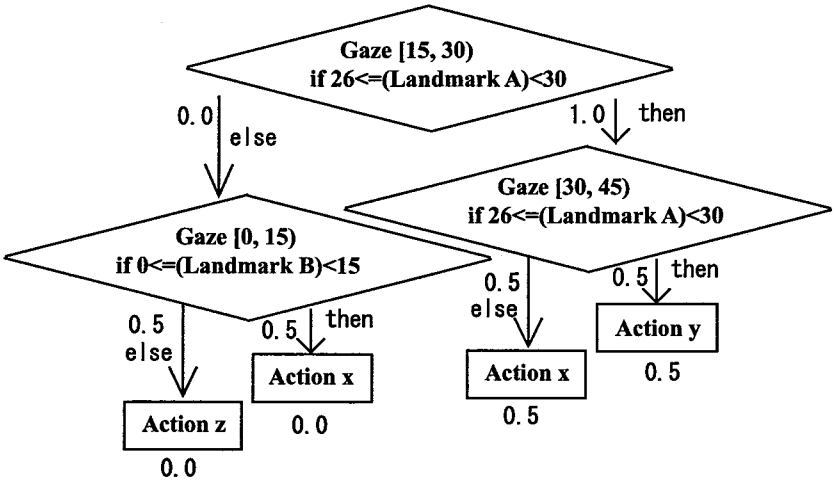
4.3.6 履歴を考慮した行動決定木の再生成

過去の観測や行動の履歴を利用することで行動決定までの観測の減少が期待される．しかし行動決定の際に決定に必要な観測しか行わない手法では，過去に十分に観測されているとは限らない．過去の観測を直接利用し，そのような場合を網羅した行動決定木を生成すると木が極端に大きく，実用に適さないと考えられる．

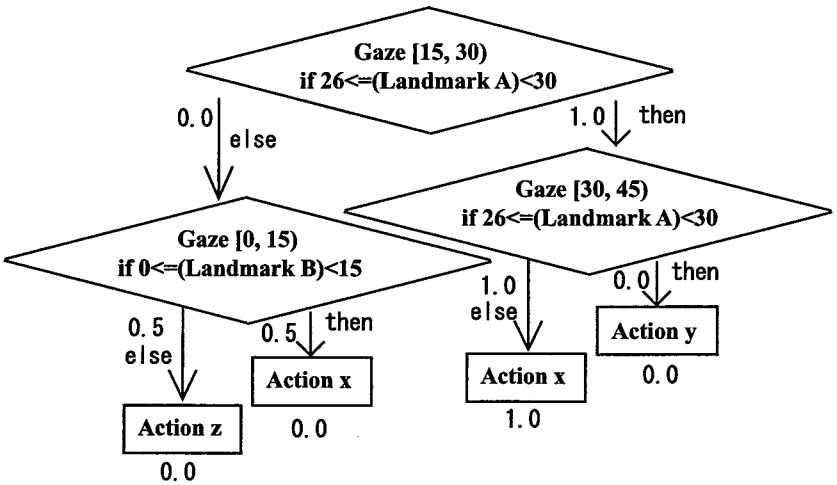
そこで過去の行動決定過程に関係なく得られる情報のみを用いることを考える．生成された行動決定木の葉に注目すると，葉はある条件下でとるべき行動を表しており，根から葉までの経路は，その条件を満たすセンサ空間を表している．このセンサ空間は同一の行



(a) Without observation



(b) Gaze [15, 30)



(c) Gaze [30, 45)

Fig.4.3 Calculation example of action probabilities

動をとる状態を表していると考えることができる．そこで以下では，この空間を同一行動状態と呼び，これを用いて行動決定木を再生成し，観測時間の短縮を図る．

行動決定木は次のように再生成する．まず，各トレーニングデータの属している同一行動状態と，一時刻前の同一行動状態を調べておく．各シーケンスの先頭については，一時刻前の同一行動状態は不明としておく．不明も一つの状態である．情報量速度は，行動に関するものではなく同一行動状態に関する情報量を計算する．

(1) 同一行動状態に関する情報量速度 i_{ik} の最も大きい i, k の組を求める．

(2) 一時刻前の同一行動状態によりトレーニングデータを2組に分けた場合に，最も情報量の大きくなる分け方を求める．同一行動状態を知るのに必要な時間は 0 であるから，この情報量を T_C で割った値を情報量速度とする．

(3) (1),(2) の情報量速度を比較し，大きい方の分割法でトレーニングデータを分ける．

(4) 同一行動状態が決定できるまで (1) から繰り返す．

ここで，行動決定シーケンスの開始状態が，トレーニングデータ中のいずれかのシーケンス開始状態に一致する場合には問題が起きないが，一致しない場合には正しく行動決定されない可能性がある．例えば，Fig. 4.4 のようにトレーニングシーケンスから行動決定木が再生成されたとする．図のトレーニングシーケンスは同一行動状態で表しており，最初の同一行動状態は p，その一時刻前は不明である．行動決定の開始時は直前の同一行動状態は不明である．同一行動状態が r である状態から行動を開始した場合には，トレーニングシーケンスでは直前の同一行動状態は p または q であるのに対し，木を使った行動決定時は「不明」となる．そのため，行動決定を誤り，x をとることとなる．

そこで，トレーニングシーケンス途中から行動を開始する可能性のある場合には，シーケンス途中のトレーニングデータについて，一時刻前の同一行動状態を不明としたデータを追加し，行動決定木を再生成する．これにより，行動決定木は一時刻前の同一行動状態が不明であっても正しく行動決定できる．

行動決定の際には，観測確率と一時刻前の各葉に到達した確率 (同一行動状態確率) から，同一行動状態確率を計算し，行動確率を求める．このとき，行動決定が可能であれば，同一行動状態が特定できなくても，それ以上，観測しない．

4.4 実験

ロボットとしては，RoboCup SONY 脚式ロボットリーグのロボット (Fig. 4.5(a)) を用いた．カメラの画角は横 52 度，縦 48 度，画素数はそれぞれ 88, 71 である．脚は各 3 自由度，首は 3 自由度 (パン，チルト，ロール) ある．ランドマークを観測する際には脚，首のロールを固定し，パン，チルト軸のみを利用した．パン軸はロボット正面に対して，-88

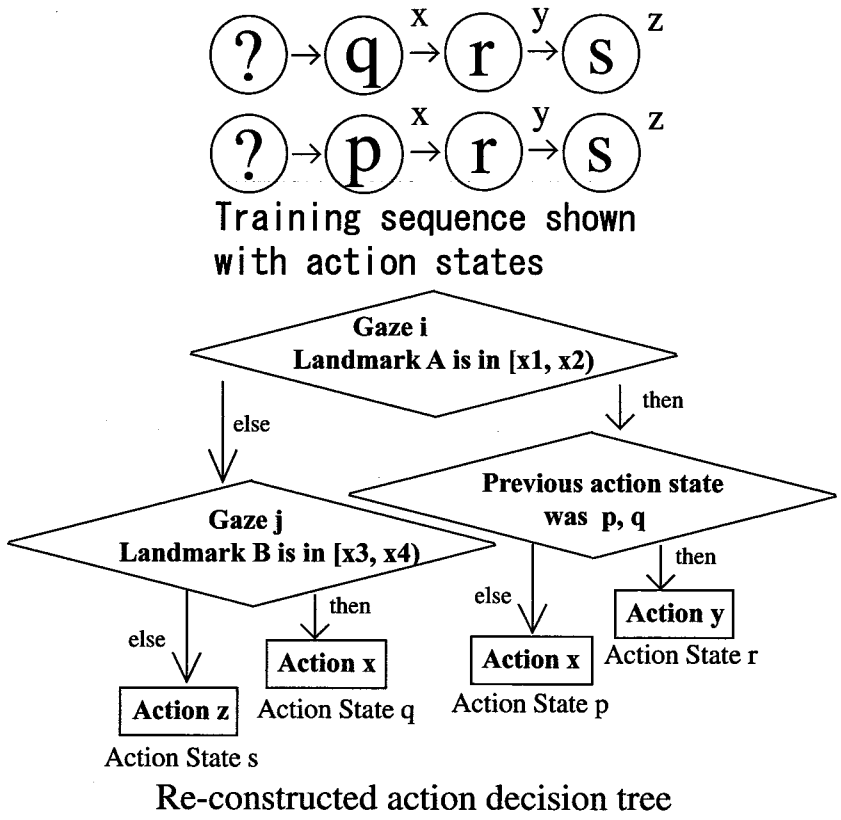
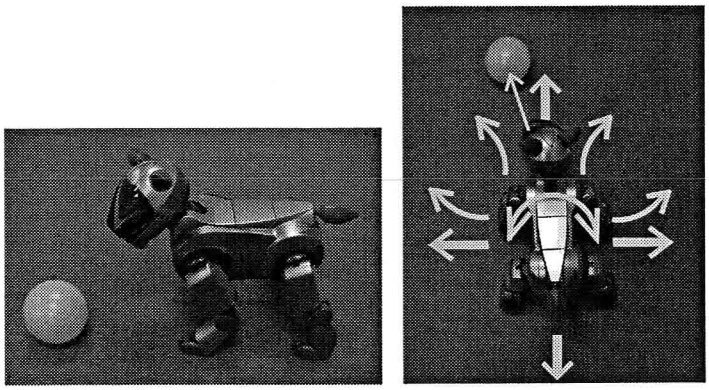


Fig.4.4 An example training sequence and re-constructed action decision tree

度から 88 度，チルト軸は，-80 度から 43 度が可動範囲である．能動的に観測する方向としては，パン軸を 44 度毎の 5 方向，チルト軸を 40 度毎の 2 方向の 10 方向にわけた．パン軸の最大角速度は $6[rad/s]$，チルト軸の最大角速度は $4[rad/s]$ である．また軸が目標値に達した後で画像が安定するまで待ってから処理をしており，カメラの方向を変えて観測する場合に少なくとも 0.36[s] かかることから，$T_C = 0.36[s]$ とした．行動としては，ボール接近，前進，後退，左(右)前進回転，左(右)その場回転，左(右)横移動，左(右)横移動回転，行動終了の 12 を用意した (Fig. 4.5(b))．

実験環境を Fig. 4.6 に示す．RoboCup SONY 脚式ロボットリーグのフィールドである．ランドマークは 6，ゴールが 2 あり，ボールが一つある．それぞれ，敵ゴール (TG), 自陣ゴール (OG), 北西ポール (NW), 北東ポール (NE), 中央西ポール (CW), 中央東ポール (CE), 南西ポール (SW), 南東ポール (SE) とする．すべてのランドマークとボールは色情報により識別される．

視覚センサとしては，各ランドマークとボールの図心座標，両ゴール (TG, OG) の画像上での (x, y) 座標が最小/最大となる座標 4 つを用いた．x,y の直積 (長方形領域) により



(a) (b)

Fig.4.5 A robot for the RoboCup SONY legged robot league (a) and prepared moving actions (b).

注視窓を構成し，トレーニングデータを分割した．これは，例えば x 座標の範囲が -10 度から 10 度にあるか否かで分割すると，y 座標に関しては指定されていないので y 軸についてカメラを回転して観測する必要があるためである．また，あまり小さな注視窓はノイズに反応することになると思われるので，注視窓の大きさに下限を 10 画素とした．

4.5 実験結果

ロボットがボールと TG を正面に観測する位置へ移動することをタスクとした．これを実現するためには場所に応じたボールへの回り込み，ボールの探索などが必要となる．ボールの位置を 3 個所，ロボットの行動開始位置を 5 個所とし，各点から目標位置姿勢までの行動を教示した．

予めセンサ値を 20 画素毎に離散化しておいた場合と，情報量を用いた離散化をした場合，情報量速度を用いて離散化した場合，さらに同一行動状態を用いて行動決定木の再生成した場合 (シーケンス途中からの行動開始への考慮なし/あり)，を比較した．教示した行動ステップ数 97 回の時点でのトレーニングデータから，それぞれの手法により得られた注視窓を Fig. 4.7〜Fig. 4.11 に示す．図中数字の単位は画素である．横軸が x 座標，縦軸が y 座標で，マス目は各観測方向にカメラが向いた際の観測可能範囲を示している．マス目の中の長方形が注視窓である．予め離散化した場合 (Fig. 4.7) と情報量による離散化 (Fig. 4.8) を比較すると，適切な離散化は均等に分けることではないことが分かる．情報量による離散化 (Fig. 4.8) と情報量速度による離散化 (Fig. 4.9) を比較すると，観測時間の考慮により注視窓のある観測方向が 6 から 3 に減少している．また注視窓を正面の方

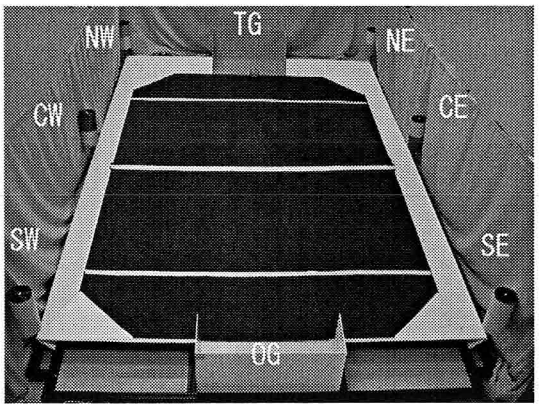


Fig.4.6 The experimental environment (the same configuration for RoboCup SONY legged robot league)

向に集中させることで，情報量の大きい注視窓ではなく，情報量速度が大きく，すでに観測した方向にある注視窓から行動を決定することにより，観測時間を減少しようとすることが分かる．履歴なし (Fig. 4.9) とあり (Fig. 4.10, Fig. 4.11) を比較すると，注視窓のある観測方向はほとんど同じであり，注視窓の傾向も似ている．

Table 4.4 に，予めセンサ値を離散化した場合 (pre-quantized)，情報量による離散化をした場合 (info. gain)，情報量速度による離散化をした場合 (info./time)，木の再生成をした場合 (re-construct)，シーケンス途中からの行動開始を考慮した木の再生成をした場合 (re-construct(M)) について，ノード数 (# of nodes)，木の最大深さ (max. depth)，葉数 (# of leaves)，観測方向数の期待値 (# of dirs)，観測時間の期待値 (time) の比較を示す．観測方向数と観測時間の期待値は観測予測なしの場合のものである．提案手法によるセンサ値の離散化が木の大きさ，観測方向数，観測時間の減少に有効であることが分かる．また情報量速度の利用により，木は少し大きくなるが，観測時間は半分以下に改善されている．一時刻前の同一行動状態を用いると，木は大きくなるが必要な観測が減少し観測時間がさらに短くなっている．シーケンス途中からの行動開始を考慮すると，履歴なしの場合とほぼ同一であるのは，再生成された木で同一行動状態が使われていないためと考えられる．Fig. 4.9 と Fig. 4.11 で注視窓がよく似ているのも同じ理由である．一時刻前の同一行動状態が使われていないのに，観測時間が減少している理由は，センサ空間のある程度まとまった同一行動状態についての決定木を生成しているためと考えられる．

Table 4.5 に，生成された木を使って実機で行動を決定した場合の，平均観測方向数 (average gaze directions) と平均観測時間 (average time) を示す．期待値とほぼ同じ値が実現されている．

センターライン上のランドマーク CE 近傍の位置から行動開始したときの，行動決定時の各葉への到達確率，同一行動状態の確率，行動確率の変化の様子を Fig. 4.12〜Fig. 4.15

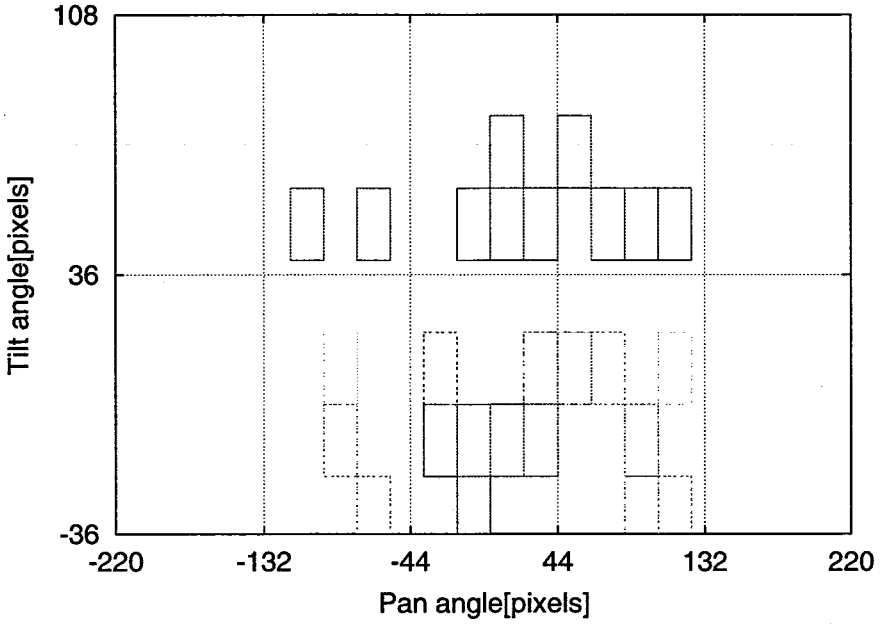


Fig.4.7 Generated attention windows with pre-quantized sensor values

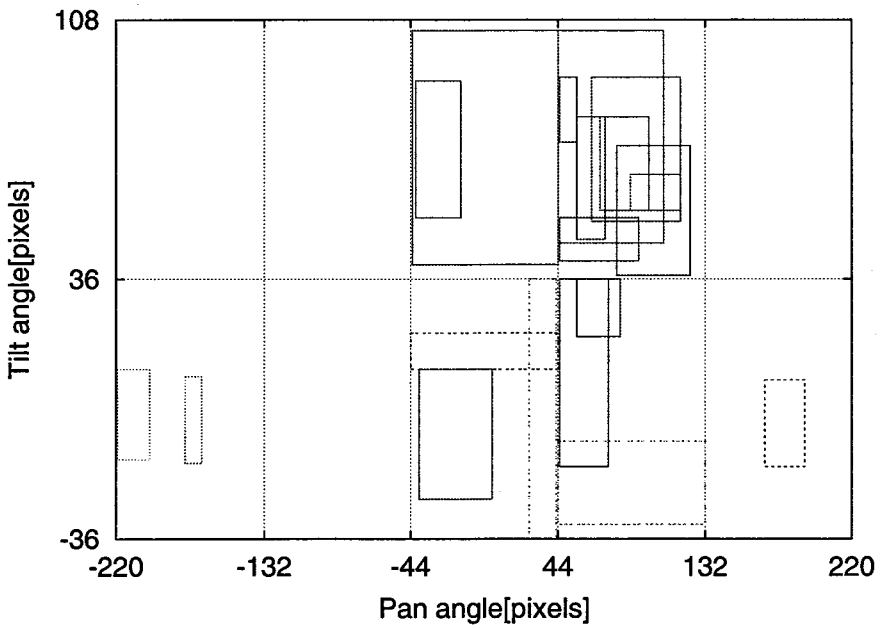


Fig.4.8 Generated attention windows with quantization by information gain

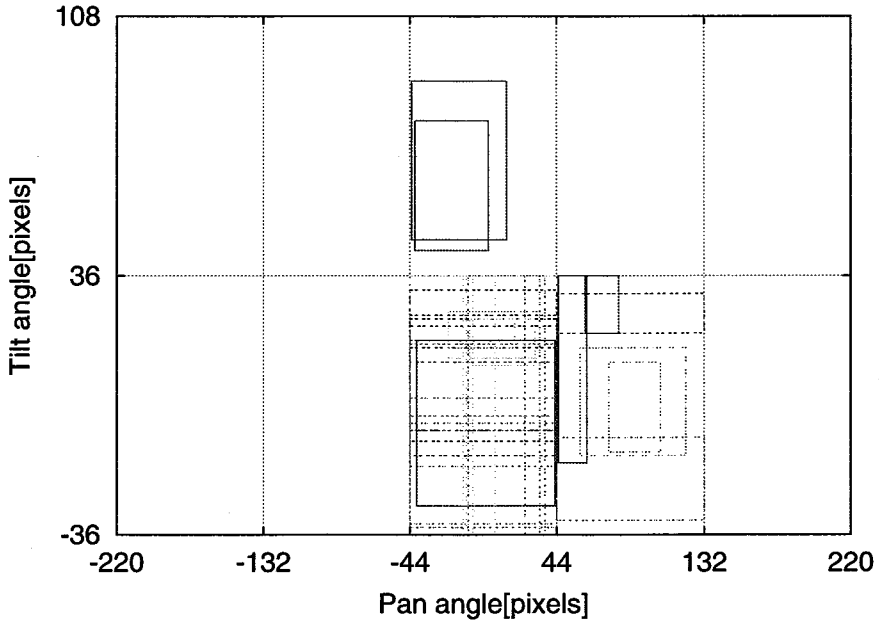


Fig.4.9 Generated attention windows with quantization by information gain per time

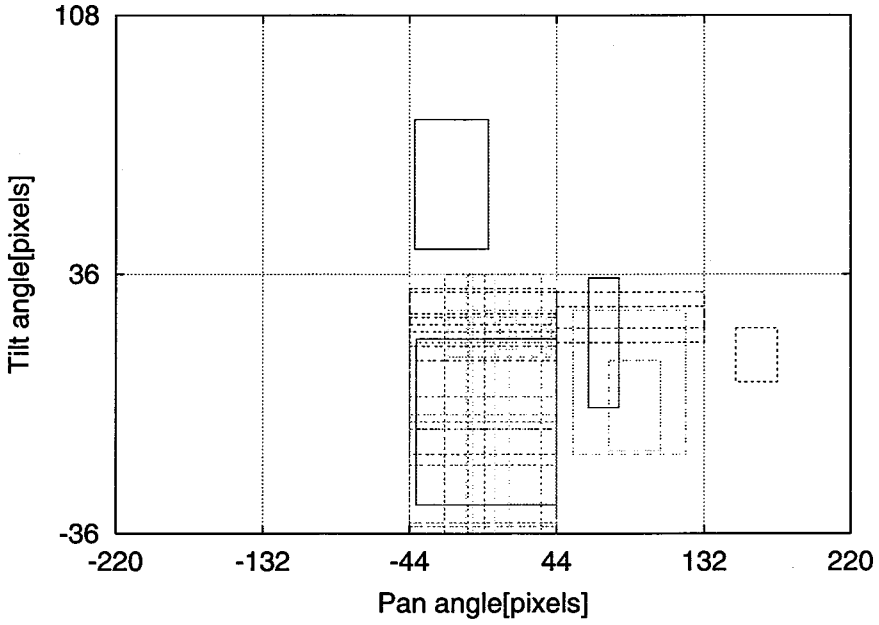


Fig.4.10 Generated attention windows of the re-constructed tree. Quantization was done by information gain per time.

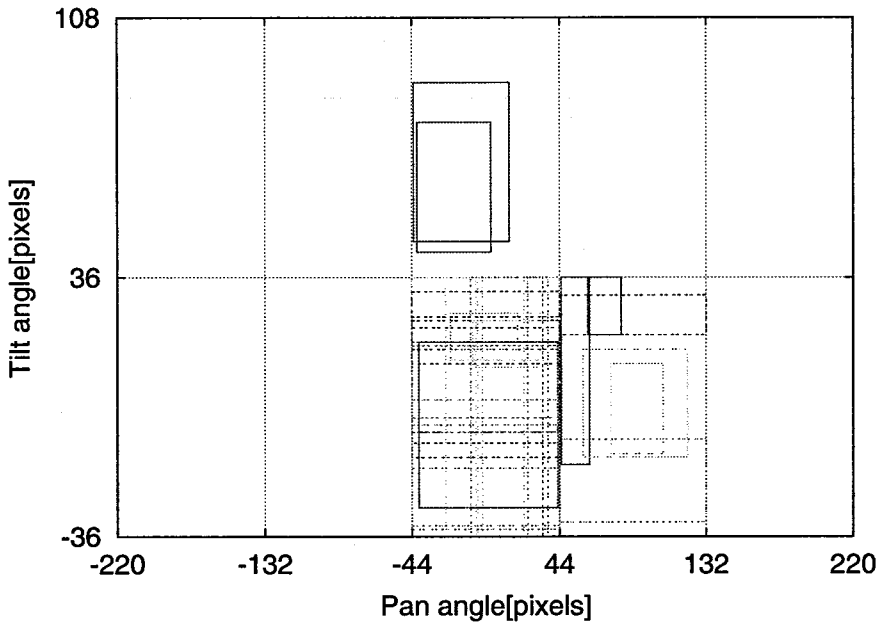


Fig.4.11 Generated attention windows of the re-constructed tree. Quantization was done by information gain per time. Action decisions started from middle of the training sequence were considered

Table 4.4 Comparison of the size of trees, expected number of gaze directions, and time to make a decision.

	# of nodes	max. depth	# of leaves	# of dirs	time[s]
pre-quantized	41	7	42	4.4	2.5
info. gain	19	7	20	3.7	2.2
info./time	29	9	30	1.5	0.84
re-construct	34	9	35	1.4	0.60
re-construct(M)	29	7	30	1.5	0.82

Table 4.5 Comparison of average number of gaze directions, and time to make a decision in experiments.

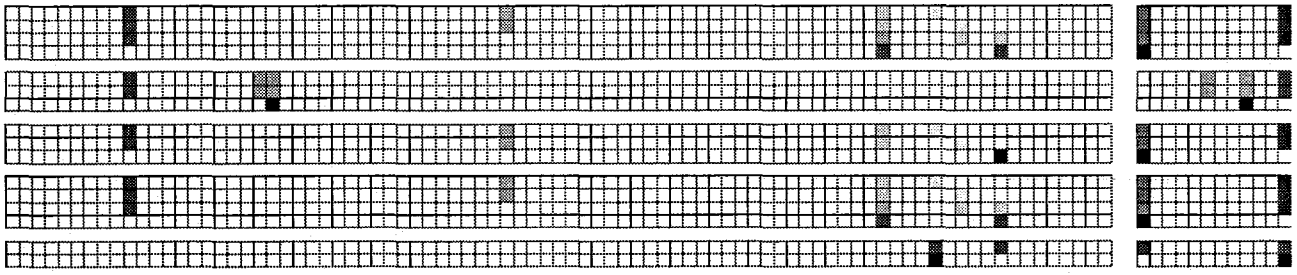
	# of data	average gaze directions	average time[s]
pre-quantized	34	3.1	3.3
info. gain	43	3.5	3.3
info./time	35	1.3	0.85
re-construct	45	1.0	0.62
re-construct(M)	42	1.4	0.97

に示す．各マスは確率を表し，白が0，黒が1である．上から下に時刻が進んでおり，観測方向が変わる毎に確率が計算され段が変わっている．空白行は行動決定し，行動を実行中であることを示す．Fig. 4.12(c) に示すように，各図の左側のマスが各葉へ到達する確率，右側のマスが各行動の確率を表す．Fig. 4.14 の中央のマスは，各同一行動状態である確率を表す．

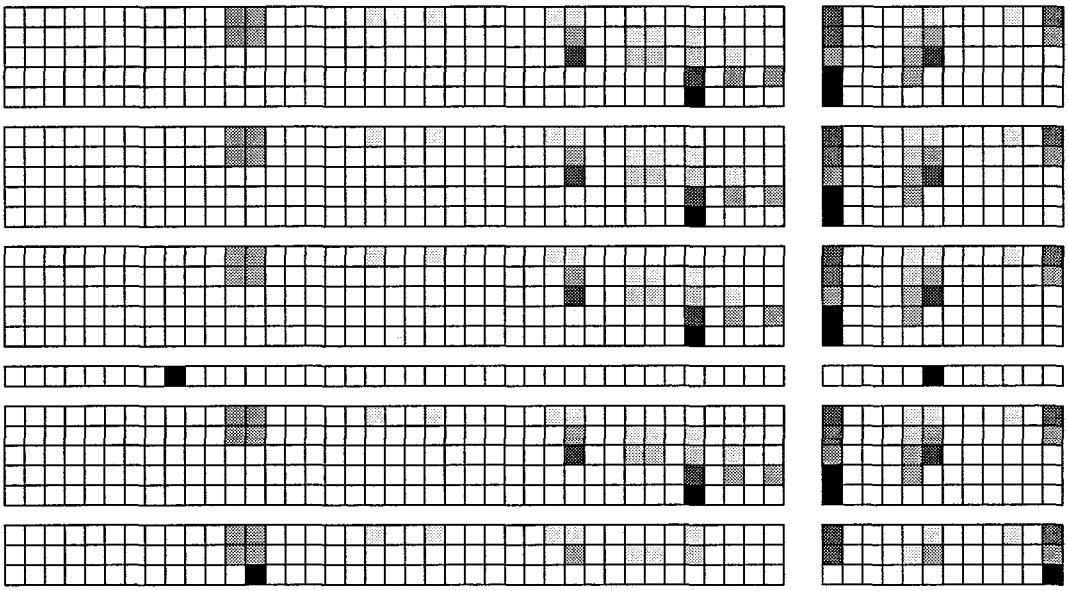
情報量による離散化では行動決定までに観測を数回繰り返していることが分かる (Fig. 4.12, Table 4.5)．また Fig. 4.12(a) を見ると，葉への到達確率が 1 に近づく (状態が識別される) 前に，行動が決定される場合のあることが分かる．情報量速度を用いた決定木により，1または2方向の観測で行動が決定されていることが分かる (Fig. 4.13，Table 4.5)．この例では情報量速度により十分に観測時間が短縮されているため，同一行動状態による木の再生成の効果は現れていない (Fig. 4.14)．

教示した行動ステップ数が 53 回の時点での情報量速度を用いた決定木の場合には，異なる行動決定がされた (Fig. 4.15)．最初に決定した行動が異なるのは，初期位置姿勢の少しの差により異なった行動が選ばれているためである．これはトレーニングデータに同じ位置で異なる姿勢の場合に別の行動を与えたものを，微妙な姿勢差で識別しようとしたためと考えられる．行動ステップ数が 97 回の場合 (Fig. 4.13) と比較してシーケンスが短いのは，ロボットがボールの前にたどり着く前に行動終了を選択したためである．これらは，トレーニングデータの増加により解消されたと考えられる．

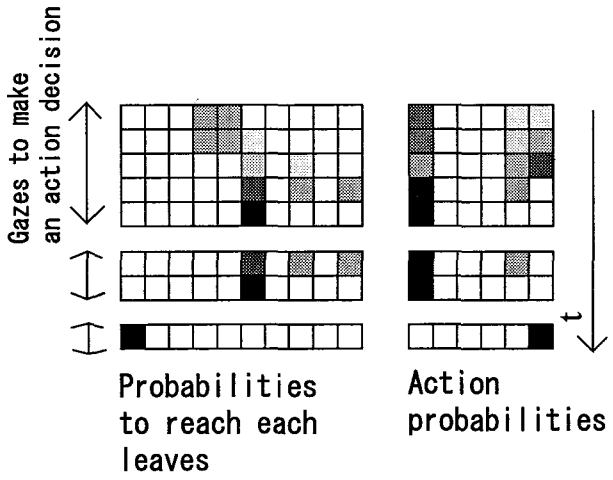
同一行動状態を用いて行動決定木を再生成した場合 (シーケンス途中からの行動開始への考慮あり) により得られた行動決定木の一部を Fig. 4.16 に示す．図中数字の単位は画素である．まずゴール TG の位置と大きさによって行動決定しようとしている．教示された条件ではボールとゴールの位置が限られているため，行動決定にボールを利用しない場合も多くみられる．またロボットの位置を決めて，ボールのボール位置を変えて違う行動を明示的に教示することは行わず，ボールとロボットの位置が毎回少しずつ異なっていたためであると考えられる．



(a) Fixed qunatization



(b) Proposed quantization



(c) Legend

Fig.4.12 An example action decision sequence by the tree constructed by information gain. The probabilities to reach each leaves (left) and the action probabilites (right) are shown. The intensity of each small box indicates the probability, white means zero and black means one.

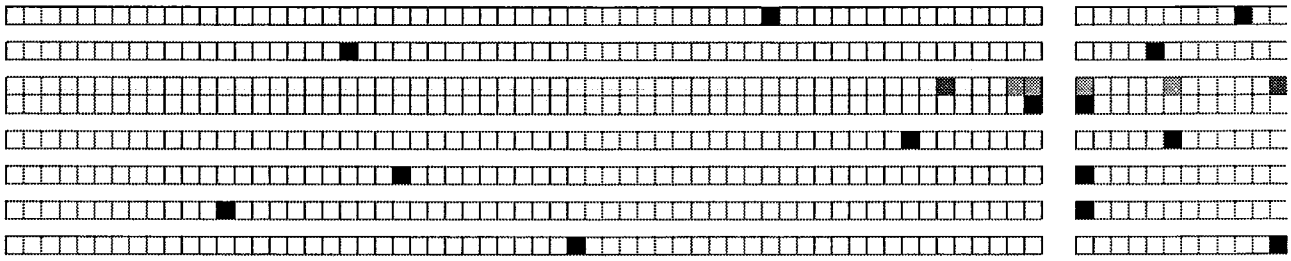
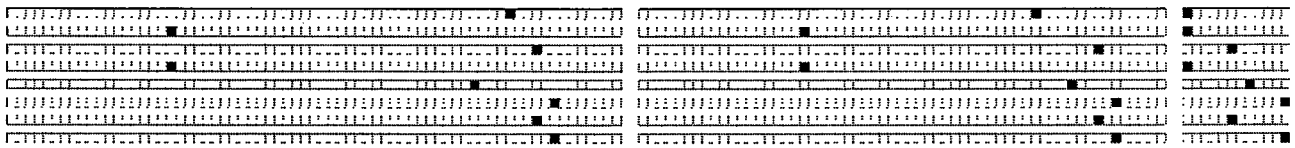
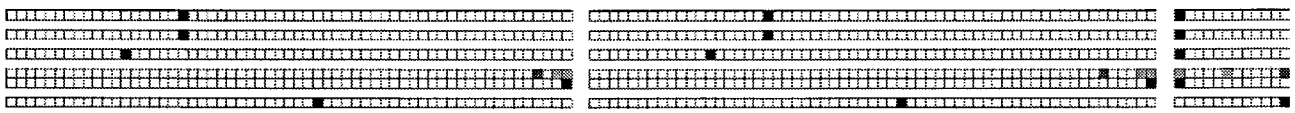


Fig.4.13 An example action decision sequence by the tree constructed by information gain per time. The probabilities to reach each leaves (left) and the action probabilites (right) are shown.



(a) Without consideration to start from middle of the training sequence



(b) With considertation

Fig.4.14 An example action decision sequence by the re-constructed tree with action state. The original tree was constructed by information gain per time. The probabilities to reach each leaves (left), the action state probabilities (center), and the action probabilities (right) are shown.

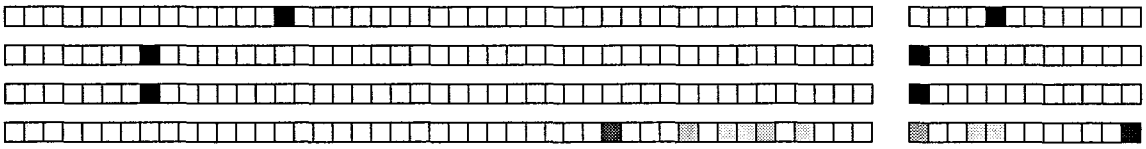


Fig.4.15 Another example action decision sequence by the tree constructed by information gain per time. We used 53 training data for this tree. The probabilities to reach each leaves (left) and the action probabilites (right) are shown.

Table 4.6 Comparison of the size of trees, expected number of gaze directions, time to make a decision, and average attention window sizes.

# of data	# of nodes	max. depth	# of leaves	# of dirs	time[s]	window size (x)	window size (y)
22	8	6	9	1.5	0.85	81	19
35	18	7	19	1.5	0.81	60	53
47	19	7	20	1.5	0.85	54	45
53	22	10	23	1.7	0.95	44	41
80	26	8	27	1.5	0.83	60	34
88	30	9	31	1.5	0.84	55	39
97	29	9	30	1.5	0.84	55	37

次に，教示データの増加に伴う行動決定木のノード数(# of nodes)，木の最大深さ (Max. depth)，葉数 (# of leaves)，観測方向数の期待値 (# of dirs)，観測時間の期待値 (Time)，注視窓の x, y 軸方向の平均の大きさ (window size x/y) の変化を Table 4.6 に，生成された注視窓の変化を Fig. 4.17，Fig. 4.18 に示す．教示データは行動決定木を用いて行動決定した際に誤った決定を行った場合に，追加することで増加した．行動決定木の生成には情報量速度を用いた．Table 4.6 を見ると，データ数が増加するにつれて，木は大きくなるが，単調増加ではない．また観測方向期待値と観測時間期待値は，データ数に対してさほど変化していない．生成された注視窓の変化 (Fig. 4.17, Fig. 4.18) を見ると，基本の窓はそのままで，誤った判断を導いた窓が修正されていったように見える．またランドマークが複数あり，行動決定にいずれを用いてもよいので，観測した方向にあるランドマークを使う傾向にあることを示していると思われる (Fig. 4.17(d))．

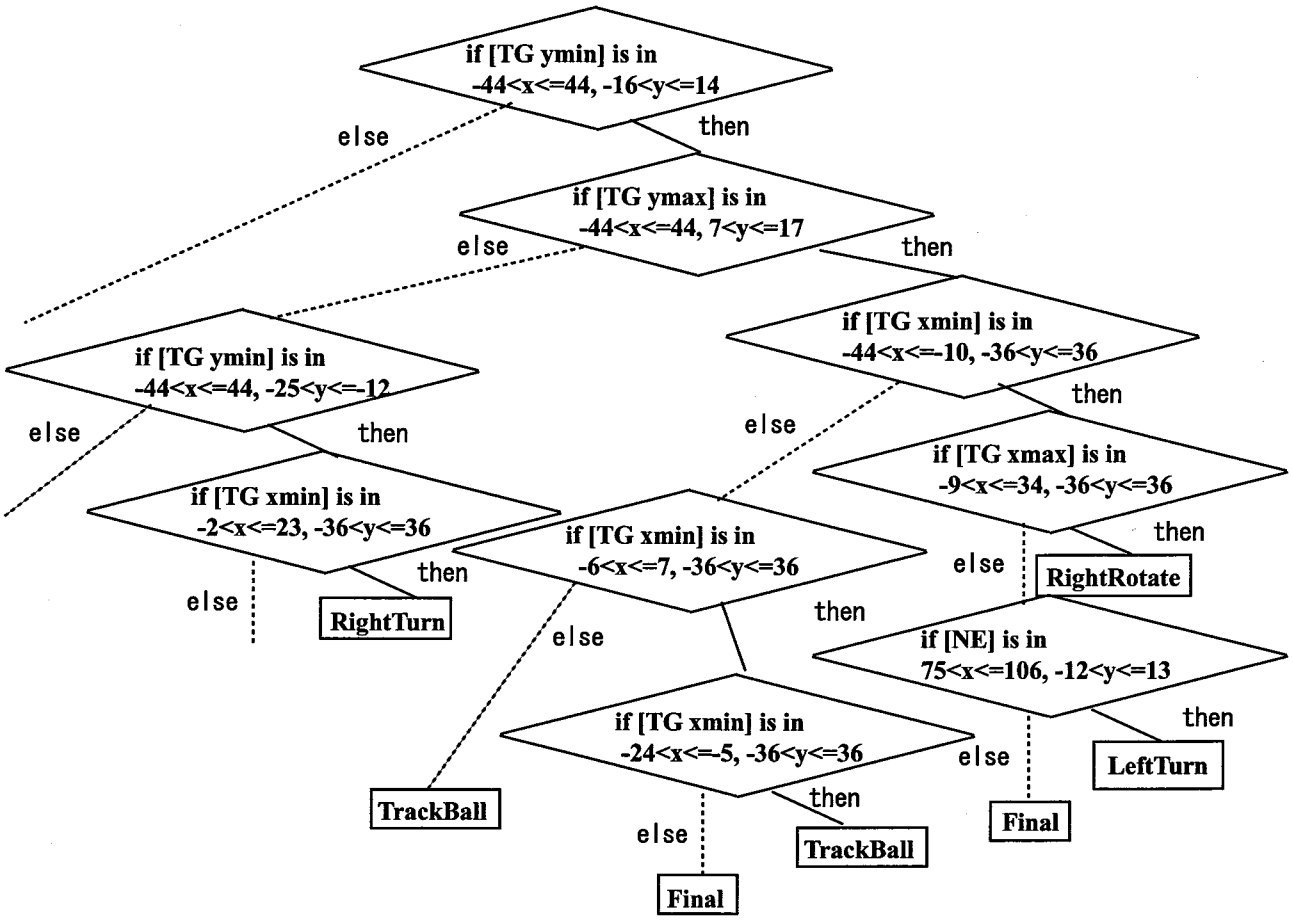
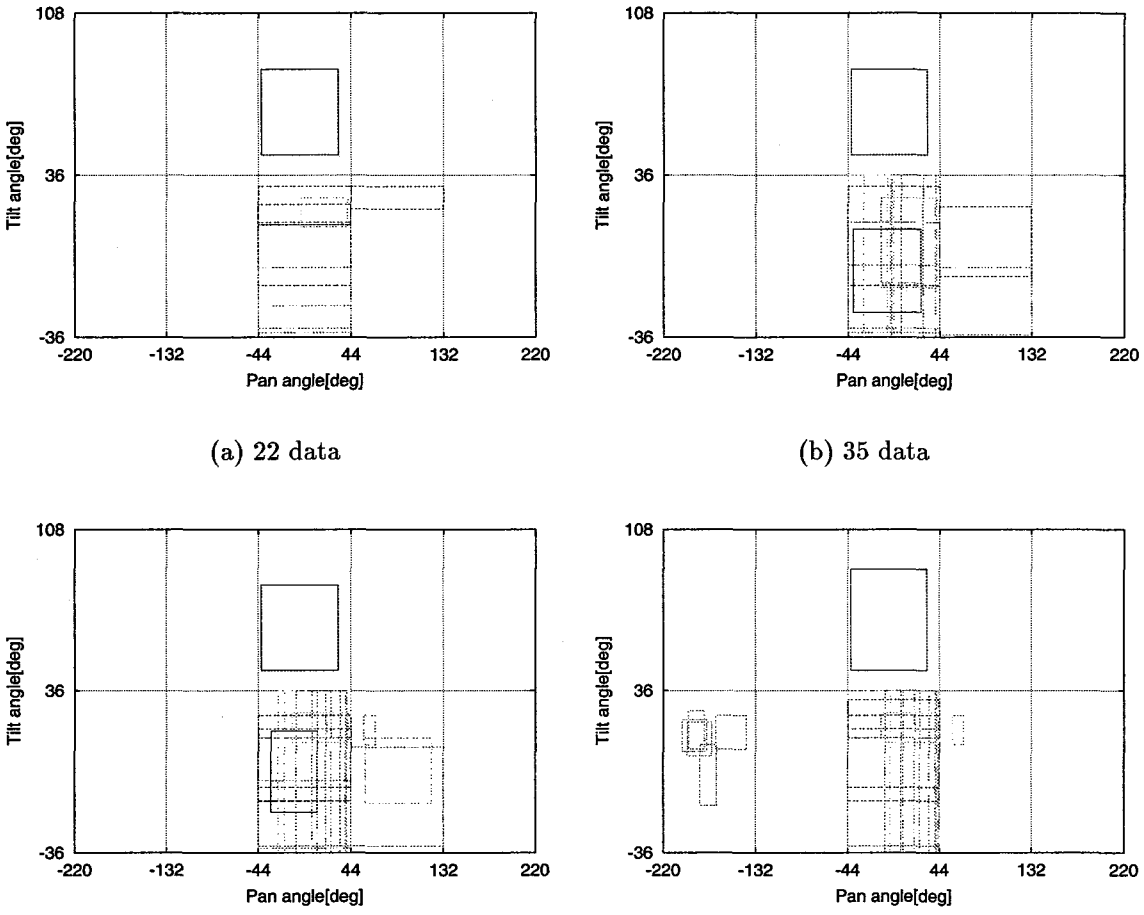


Fig.4.16 A part of the action decision tree constructed by the proposed method



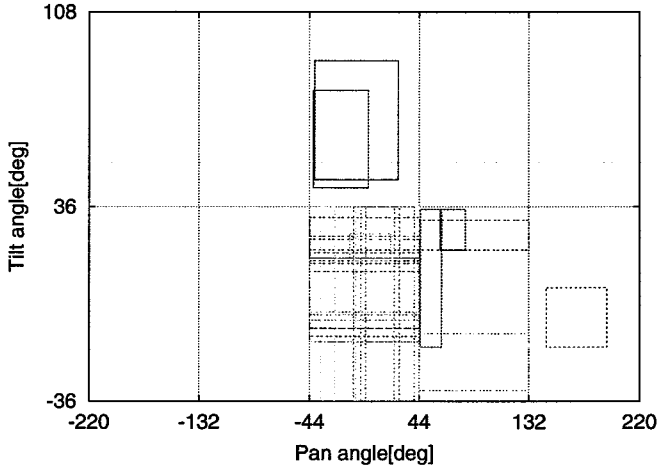
(a) 22 data

(b) 35 data

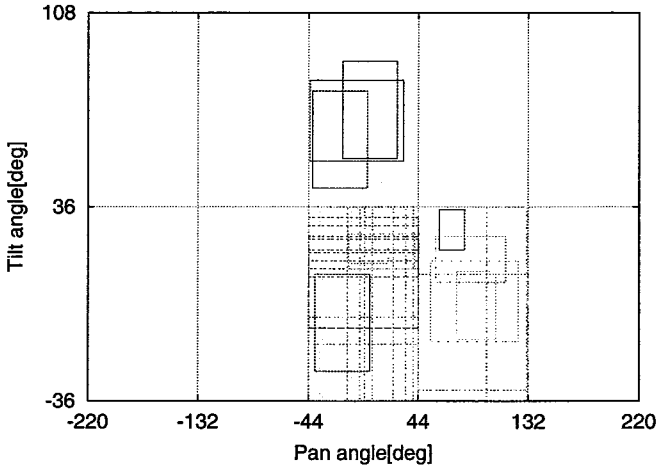
(c) 47 data

(d) 53 data

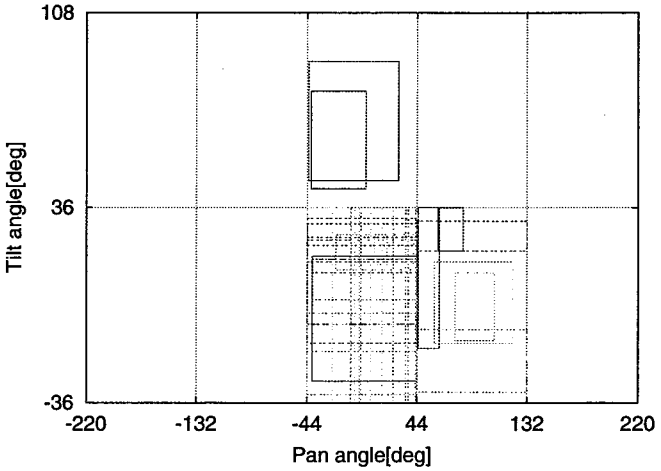
Fig.4.17 Changes of attention windows along with the increase of data(1/2)



(a) 80 data



(b) 88 data



(c) 97 data

Fig.4.18 Change of attention windows along with the increase of data(2/2)

4.6 結言

自己位置の同定ではなく行動決定のための，視覚センサと観測時間を考慮した，効率的な観測をする行動決定法を提案した．実験により提案手法が，適切な離散化による行動決定木の圧縮と，観測時間の短縮に有効であることを確認した．

行動決定木により決定された行動が期待したものとはならない場合がある．一つは，行動確率の閾値が低い場合である．これは観測が少ないときに行動確率の高い行動と，十分な観測を得た後に確率が高い行動が必ずしも一致しないためである．これには行動確率について信頼度の導入による解決が考えられる．もう一つは，トレーニングデータが不足している場合である．情報量速度の高い観測を優先するため，データが少ないときには，トレーニングデータに含まれたノイズを行動決定に利用した決定木を生成してしまうことがある．行動決定木のトレーニングデータに対する汎化能力は，トレーニングデータに含まれるランドマーク方位間の内挿および，それらから注視窓内外を決める閾値 (ここでは方位間の中点あるいは，カメラの画角) までの外挿が期待出来る．汎化の程度は，センサ空間のタスクとの親和性に依存する．タスクに適当なセンサ空間，必要なトレーニングデータの見積もり，トレーニングデータの追加時の部分的な決定木の修正の実現と観測時間への影響などは今後の研究課題である．

第5章　歩行ロボットのための情報量に基づく注視制御

5.1　緒言

第3章，第4章で，情報量を基準とした行動決定木による，効率的な観測手法を提案した．ロボットのタスク遂行のための行動決定に幾何学的な自己位置表現が必ずしも必要ではないことを指摘し，情報量を基に生成した行動決定木を用いた，明示的な自己位置表現を使わない，行動決定のための効率的な観測手法を提案した．現在と過去の観測から観測確率を求め，行動決定木を基に行動確率を計算し，行動決定する．しかし，移動ロボットが観測，行動決定する際には静止していることを前提としていた．より効率のよい観測を実現するためには，静止状態での観測だけではなく，移動中の観測も利用できる必要がある．

移動中の観測の利用と行動決定に関しては，

1) 静止時と比較してセンサに揺れなどの外乱が加わる，

2) 視覚情報の処理に比較的長い時間がかかり，処理結果が利用可能になったときには無視できないほど移動している場合や，複数方向の観測 (首振り) の統合の際に観測間で移動が無視できない場合がある，

3) 観測の曖昧さから行動決定に曖昧さが残っている場合には，衝突などの問題を避けるため，移動速度を低下あるいは，静止観測の必要がある，

等の問題がある．情報量を基に生成された行動決定木を用いた行動決定においては，1) 歩行による揺れの補正，2) 移動中の複数観測の統合，3) 適切な行動決定の指標，の解決が必要となる．

本章では，行動決定の曖昧さの指標として，行動確率と共に，観測により将来得られると期待される情報量を曖昧さの指標として用い，行動確率あるいは情報量期待値が閾値を満足しない場合には，移動せずに観測する戦略を用いて上記 3) の問題に対処する．また，歩行による揺れや，視点移動の影響を軽減する画像補正メカニズムを提案し，上記 1)，2) の問題に対処する．画像列のみから補正値を計算し，それにより視覚センサによる観測から，静止時にある観測が得られる確率を計算し，複数観測の統合と揺れの補正を実現する．

以下では，まず歩行中の観測を想定した場合の，情報量に基づいた行動決定木による注視制御と行動決定の問題点を考察する．次に観測による情報量の期待値を行動決定の信頼度として用いることを提案し，決定木による行動決定法を述べる．次に，移動による観測のずれと，歩行による揺れの補正値の計算法を提案する．そして，小型4脚ロボットを用いた実験結果を示し，結論を述べる．

5.2 歩行中の観測を考慮した注視制御・行動決定問題

多くの場合，行動決定には観測情報すべてが常に必要ではなく，選択的に観測することで，行動決定を効率化できる．一方，効率の良い観測戦略を得るためには，すべての可能な観測の結果を得て観測の効率を比較する必要がある．たとえば，視野角が限られているが，能動的に観測方向を変えることが出来る視覚センサを持つ移動ロボットの場合には，可能な観測方向すべてについて観測して得られた情報が必要である．

各状況で，可能な観測方向すべてについて観測して得られた情報と，その状況でとるべき行動が与えられたとする．このとき行動に関する情報量を各注視対象の観測について計算することができる．情報量が高い注視対象は，行動について曖昧さの減少が大きく，行動決定に重要である．従って行動決定できるまで情報量の高い順に注視対象と観測方向を変えていくことが，効率のよい観測戦略となる．この観測戦略は，得られた観測により枝分かれする行動決定木の形で表すことが出来る (第3章，第4章)．

5.2.1 必要となる画像補正

静止時に観測して行動決定する場合には，各状況での観測情報とそれに対応する行動を，教示などにより与えることが出来る．しかし，歩行中に観測する場合には，1) 歩行による揺れ，2) 視点の移動が問題となる．揺れにより同じ状況でも観測が異なるため，観測戦略はそれを考慮したものでなければならない．各状況での歩行中の揺れにより異なる観測を十分用意することができれば，情報量による観測戦略は揺れを考慮したものとなると期待される．しかし，各状況毎に揺れによる影響を含んだ観測データを十分に用意することは現実的でない．

また視点移動も問題となる．観測戦略を得るために観測情報を収集する際には，同じ状況で視線方向を変えて観測しなければならない．そのためには，歩行を一時停止し，静止して視線方向を変え観測を繰り返すことが出来ればよい．しかし，動歩行の場合には，その場で姿勢を保って静止することは出来ない．一方，観測方向を変えて何度も同じ経路を通るよう移動し，観測を行なうことも考えられるが，現実的でない．また，行動決定前に視点が移動してしまう可能性があることも，考慮しなければならない．

観測情報の収集の問題は，静止して視線方向を変えた観測を基に歩行中の観測戦略を決

定できれば，解決する．そこで，1) 観測情報を歩行の揺れについて補正し，静止時の観測に変換すること，2) 視点移動の問題については，観測時から現在までの移動分の補正を観測情報に施すことで，静止時の観測から得た行動決定木を歩行中の観測に利用することを考える．ここでは，1)，2) のために，歩行中の画像列から計算された補正値を用いることを提案する．補正値の計算には画像列のみを用い，他のセンサは用いない．補正値を用いて，静止時にある観測が得られる確率 (観測確率) を計算することで，複数観測の統合と揺れの補正を実現する．行動決定木から観測確率を用いて，各行動のとるべき確率 (行動確率) を計算し，行動決定する．短い周期で観測確率の計算と行動決定を繰り返し，必要に応じて視線方向を変えた歩行中の観測や，静止しての観測を行う．

5.2.2 歩行中の観測を考慮した意思決定

歩行中の観測には，必要に応じて移動を制限し観測を行なうための適切な指標が必要である．行動確率そのものを指標として用いることも考えられるが，適当ではない．なぜなら，観測が少ないときにも，ある行動の確率が高くなることがあり，誤った行動決定を避けるためには行動確率の閾値を高くしなければならない．

静止，観測，行動を一定周期で繰り返す行動決定の場合に，閾値が高いと行動決定までの観測時間は長くなるが，それほど大きな問題ではない．また観測順については，同一場所での静止時の観測のみを用いることから，固定した順序でよい．行動決定木の上位ほど観測による情報量が大きいことから，木の上位の方向から観測することで観測時間が短くなる．

一方，歩行中に観測する場合に行動確率の閾値を高くすると，頻繁に静止し観測することになり，従来の静止時観測と変わらない．ここで行動確率についての信頼度の指標があれば，行動確率の閾値を下げることが出来ると考えられる．そこで，行動確率と共に，観測により将来得られると期待される情報量 (情報量期待値) を行動の曖昧さの指標 (信頼度) として用いることを提案する．行動確率あるいは情報量期待値が閾値を満足しない場合には，移動せず観測する戦略をとる．

また移動中の観測など，観測場所の異なる観測を利用する場合には，移動により以前の観測は次第に曖昧さが増大し，再観測が必要となるため，固定した観測順では対応できない．また観測確率のみからは，観測の重要度が比較できない．情報量期待値は観測の曖昧さと重要度を反映している．そこで情報量期待値の大きい (信頼度の低い) 方向から順に観測する．

5.3 情報量による注視選択と歩行への対応

ロボットや環境，与えられるデータ等に関して以下のように設定する．

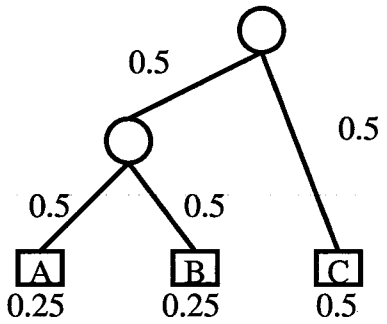


Fig.5.1 An example action decision tree.

1. ロボットは視覚センサを持つが視野角が限られており，受動的な観測のみでは行動決定に必要な情報が得られるとは限らない．
2. 環境中にランドマークが配置されており，視覚センサの観測方向を変えることにより視野を拡大し，行動決定に十分な情報が得られる．
3. 教示などにより，視野を拡大した際に得られる視覚センサの値と，その際にとるべき行動が与えられる．これをトレーニングデータと呼ぶ．

トレーニングデータを用意する際には，同じ状況で注視対象による情報量の差異を知るために，ロボットは静止して観測する．また行動決定に必要な観測情報は一定ではなく，状況に応じて変化する．歩行による画像の乱れを補正するためのデータは別に与えられる．

5.3.1 行動決定

行動を決定するには，まず，モデルを使い各注視窓にランドマークが入る確率を，現在の観測あるいは，過去の観測からの予測により計算する．いずれも得られない場合には，各注視窓に入る確率 (観測確率) は 0.5 とする．つぎに，観測確率を用いて行動決定木の各葉への到達確率を計算する．同じ行動を示す複数の葉への到達確率の和を，その行動をとるべき確率とする．行動が決定できるまで観測と確率分布の更新を繰り返す．簡単な行動決定法としては，特定の行動の確率が閾値を越えたとき，その行動をとることが考えられる．閾値を下げることで行動決定までの観測が少なくなるが，行動決定を誤る可能性が高くなる問題がある．たとえば，行動決定木が **Fig.** 5.1 である場合には，観測前に行動 C の確率は 0.5 と高い．このため，行動確率の閾値は少なくとも 0.5 を越える必要がある．

そこで，観測による情報量の期待値を信頼度の指標として用いることで，閾値を下げることを考える．観測による情報量の期待値の総和は，

$$-\sum_{all\,node}\left(\frac{n_{node}}{n}\right)I_{node}\{p\log_2 p+(1-p)\log_2(1-p)\}, \tag{5.1}$$

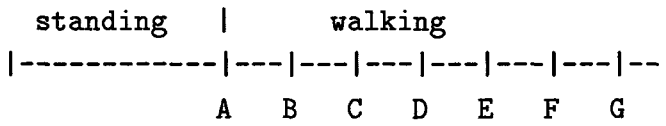


Fig.5.2 A time sequence example. At time A, it starts walking. A, B, C,... are the beginning of a walking period.

である．ただし，n_{node} は，$node$ に到達するトレーニングデータ中のデータ数で，I_{node} は行動決定木生成時に計算された情報量，p はランドマークがそのノードの注視窓内に観測される確率である．p のエントロピー $-\{p\log_2 p+(1-p)\log_2(1-p)\}$ は，0 から 1 の値をとり，最も観測が曖昧な $p=0.5$ の時，1 となる．

観測による情報量の期待値の総和が，閾値より小さくなり，ある行動の確率が閾値を越えたとき，その行動をとると決定する．いずれかを満たさないときには，ロボットは静止し，閾値が満たされるまで，情報量期待値

$$-\sum_{somedirection}(\frac{n_{node}}{n})I_{node}\{p\log_2 p+(1-p)\log_2(1-p)\} \tag{5.2}$$

の最も大きな方向の観測と，観測，行動確率の計算を繰り返す．行動決定がされた行動を実行中も，最も情報量期待値の大きい方向の観測，観測確率と行動確率の計算，行動決定を繰り返し行い，閾値を満たさなくなると静止して観測し，行動決定する．

5.3.2 画像の補正法

静止状態で観測した画像を用いて生成した行動決定木により，移動中の行動決定を行うには，何らかの補正が必要となる．一つは移動に伴う画像の変化の補正であり，車輪式，脚式に共通に必要である．もう一つは，歩行による揺れの補正である．

この 2 つを補正するため，使用する歩行と観測方向それぞれについて補正値を歩行中の画像列 (**Fig.** 5.2) から計算する．まず，ロボットは静止しており，A の時刻で歩行を開始し，B の時刻が第 2 歩行周期の，C, D, E,..., の時刻は，それ以降の歩行周期の開始時刻である．

A, B, C,.. の時刻を t_A, t_B, t_C,.. と表記し，時刻 t の画像を U_t とする．歩行の開始初期では，歩行の位相が同じ時刻の画像 U_{t_A+i} と U_{t_B+i} の差に，歩行初期の揺れと移動の両方の影響が含まれる．歩行が定常状態に入ったときには，歩行の位相が同じ時刻の画像 U_{t_F+i} と U_{t_G+i} の差は，移動による影響のみが含まれる．ここで画像の差は，画像上各点の明度差 (カラーの場合には 3 色それぞれの差) の 2 乗和，

$$D(i,j,\Delta x,\Delta y)=\sum_{com}\{u_i(x,y)-u_j(x-\Delta x,y-\Delta y)\}^2/S, \tag{5.3}$$

で定義する．ここで，$\triangle x, \triangle y$ は画像のずれ，com と $S(S \neq 0)$ は，それぞれ比較する画像の重なる領域とその面積，$u_i(x,y)$ は，画像 U_i 上の点 (x,y) の明度値である．

まず移動による画像変化の補正値を求める．画像の差 $D(t,\ t+1,\ \triangle x,\ \triangle y)$ が最小となる $(\triangle x(t), \triangle y(t))^T$ を，画像座標上でのカメラの動きとし，各時刻 t について求める．そして，歩行が定常状態である時刻 t について，一歩行周期のカメラの動き，

$$\sum_{j=t}^{t+W} (\triangle x(j), \triangle y(j))^T, \tag{5.4}$$

の平均 $\hat{\boldsymbol{x}}_L$ と 分散 $\boldsymbol{\sigma}_L^2$ を求める．ここで，W は歩行周期である．$\hat{\boldsymbol{x}}_L$ と $\boldsymbol{\sigma}_L^2$ は，移動による画像座標上での変化の平均と分散を表す．

次に，揺れによる画像変化の補正値を求めるが，静止時の画像と歩行時の画像の差には，揺れと移動の両方の影響が含まれている．そこで移動による画像変化分を引き，時刻 t，歩行位相 θ での揺れによる画像変化分を求めると，

$$\boldsymbol{x}_s(t) = \sum_{j=0}^{t} (\triangle x(j), \triangle y(j))^T - \frac{(t-t_A)\hat{\boldsymbol{x}}_L}{W}, \tag{5.5}$$

となる．t_A は歩行開始時刻を表す．第1項は静止時から時刻 t までの画像変化を，第2項は移動分の補正を表す．$\boldsymbol{x}_s(t)$ を用いて，各歩行位相 θ 毎に，画像の変化の平均 $\hat{\boldsymbol{x}}_S(\theta)$ と分散 $\boldsymbol{\sigma}_S^2(\theta)$ を計算する．これには歩行初期の揺れの影響を含まれる．

観測結果と $\hat{\boldsymbol{x}}_L, \boldsymbol{\sigma}_L^2, \hat{\boldsymbol{x}}_S(\theta), \boldsymbol{\sigma}_S^2(\theta)$ から観測確率を計算する．ランドマーク i が画像上の位置 $\boldsymbol{x}$ に時刻 t_1 に観測されたとき，時刻 t に観測が期待される画像上の位置と偏差は，

$$\bar{\boldsymbol{x}}(t) = \boldsymbol{x} - \hat{\boldsymbol{x}}_S(\theta_{t_1}) + \frac{(t-t_1)\hat{\boldsymbol{x}}_L}{W}, \tag{5.6}$$

$$\bar{\boldsymbol{\sigma}}(t) = \sqrt{\boldsymbol{\sigma}_S^2(\theta_{t_1}) + \frac{(t-t_1)\boldsymbol{\sigma}_L^2}{W}}, \tag{5.7}$$

となる．ここで，θ_{t_1} は時刻 t_1 での歩行位相である．k を正の定数として，$\bar{\boldsymbol{x}}(t) - k\bar{\boldsymbol{\sigma}}(t)$ と $\bar{\boldsymbol{x}}(t) + k\bar{\boldsymbol{\sigma}}(t)$ がつくる領域の面積と，この領域が注視窓を覆う面積の比を観測確率とする．

注視窓を確認したランドマーク以外についても，観測されたランドマークについては観測結果を記憶しておく．ロボットの移動にともないランドマークの観測される画像上の位置も変化し，観測方向をまたぐことがある．これに対応する簡便な方法は，それぞれの観測方向により得られる画像を統合した一枚のパノラマ画像上の座標で，ランドマークの位置を表すことである．以下の実験では，この方法を用いている．

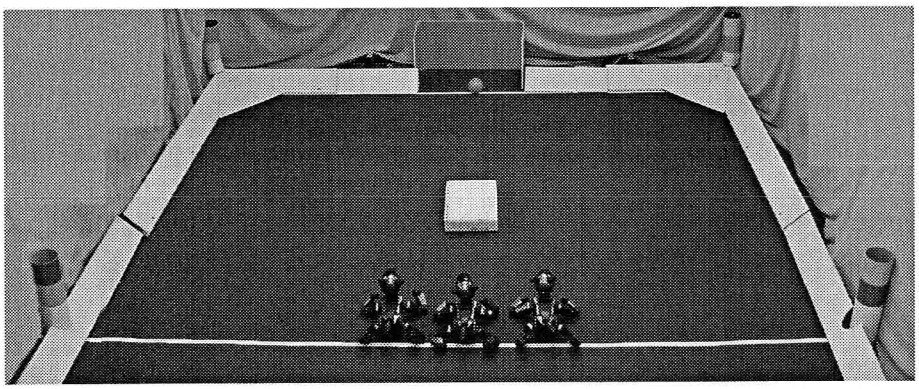


Fig.5.3 Experimental field (half of the RoboCup 2002 SONY legged robot league). Robot started from three positions on the center line.

5.4 実験

5.4.1 ロボットと実験環境

ロボットとしては，第4章の実験で用いたRoboCup SONY 4足ロボットリーグのロボットを使用した．実験環境は，第4章の実験で用いたRoboCup SONY 4足ロボットリーグ2001のフィールドに対し，長さで1.5倍に拡大されたRoboCup SONY 4足ロボットリーグ2002のフィールド(**Fig.** 5.3)を使用した．ランドマークは6，ゴールが2あり，ボールが一つあるのは共通である．それぞれ，敵ゴール (TG), 自陣ゴール (OG), 北西ポール (NW), 北東ポール (NE), 中央西ポール (CW), 中央東ポール (CE), 南西ポール (SW), 南東ポール (SE) とする．すべてのランドマークとボールは色情報により識別される．

ロボットのカメラの画角は横58度，縦48度，画素数はそれぞれ88, 71である．脚は各3自由度，首は3自由度(パン，チルト，ロール)ある．パン軸はロボット正面に対して，-88度から88度，チルト軸は，-80度から43度が可動範囲である．カメラのフレームレートは40[ms]である．以降では，時間については40[ms]を単位とする．また観測行動決定もフレームレートで行う．行動としては，前進，左前進，右前進の3つを用意した．これらの行動は，トロット歩容を基にしており，行動を用意する際には画像の揺れが少なくなるような工夫は行っていない．

視覚センサとしては，各ランドマークとボールの図心座標，両ゴール(TG, OG)の画像上での (x, y) 座標が最小/最大となる座標4つを用いた．x,y の直積(長方形領域)により注視窓を構成し，トレーニングデータを分割した．観測方向としては，パン軸は44度毎の5方向に，チルト軸は21度毎の5方向に分けた．これはカメラの揺れに関わらず特定の方向を観測したときに，必ず行動決定木生成時に仮定した範囲が観測可能とするためである．

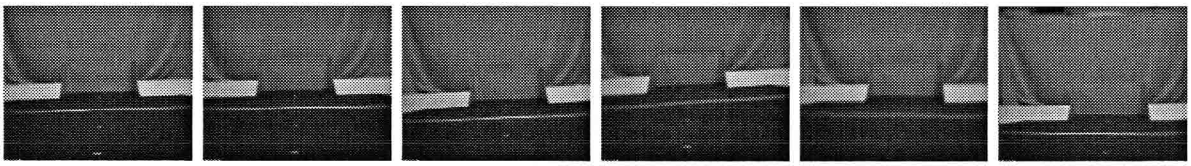


Fig.5.4 The images of the camera while the robot is moving forward. The images are taken every 80[ms].

5.4.2 画像の補正実験

Fig. 5.4 にロボットが前進行動をとり，カメラが前方を向いているときの 80[ms] 毎のカメラの画像を示す．歩行周期は 600[ms] である．**Fig.** 5.5 に，前進について 3 回の画像の変化量 $\triangle y$ を示す．一周期に一度大きく画像が乱れていることが分かる．**Fig.** 5.6, **Fig.** 5.7, **Fig.** 5.8 に，前進，右前進，左前進それぞれについて，前方を観測しているときの揺れによる変化量の平均 $\hat{\boldsymbol{x}}_S$ と偏差 $\boldsymbol{\sigma}_S$ を示す．横軸は歩行の位相である．**Fig.** 5.9, **Fig.** 5.10 に，カメラを前方に向け，ロボットがゴールを正面に見るフィールド中央の位置から，静止，前進 (25[×40ms] から 115[×40ms])，静止 (132[×40ms] まで) の行動を行ったときの，画像上でのゴールの重心位置について，補正なし，提案する揺れについての補正による中心値，歩行開始時に期待される重心位置の変化を示す．補正なしの場合には y 座標に大きな揺れが見られるのが，補正を行うことで，振幅が概ね減少している．揺れのピークで部分的に振幅が増加している部分もあるが，これは **Fig.** 5.4 に示したように，歩行による揺れに鋭いピークがあり，またピークの位相に振れがあるためである．歩行開始時と終了時に，滑りのためずれが生じているが，移動開始時に期待される重心位置の変化を表す直線と，歩行中の移動による平均的な重心位置の変化と傾きは一致していることが分かる．歩行開始時に y の補正後のずれが大きいのは，静止から定常歩行への移行が瞬時には起きないためである．また x の変化が前進から静止へ遷移する時に大きくなっているのは，トロットによる前進から静止への滑らかな遷移行動を用意していないためである．必要に応じて，滑らかな歩行遷移あるいは，歩行遷移に関する補正値を用意する必要がある．時刻 0 から 131 [×40ms] の観測から提案する補正により，期待される範囲 $\bar{\boldsymbol{x}}(t) - \bar{k\boldsymbol{\sigma}}(t)$ と $\bar{\boldsymbol{x}}(t) + \bar{k\boldsymbol{\sigma}}(t)$ に，時刻 132 [×40ms] の値が含まれる確率は，k が 1.0 のとき 12%，k が 2.0 のとき 40%，k が 2.5 のとき 73%，k が 3.0 のとき 86% となった．前進行動について複数回の試行により，k が 2.5 のとき，70 から 80% となったので，以下では $k = 2.5$ を用いた．

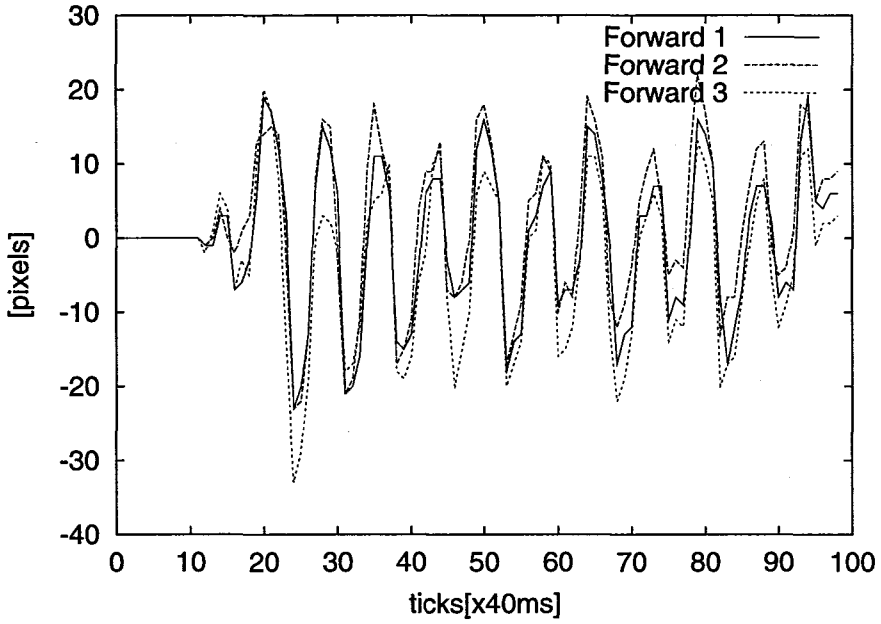


Fig.5.5 The $\triangle y$ of forward motion.

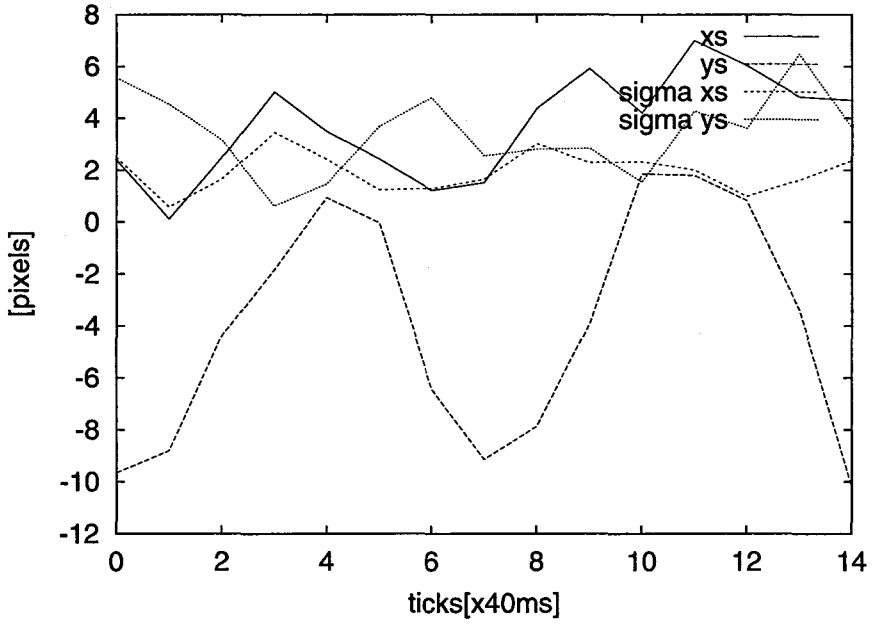


Fig.5.6 The $\hat{\boldsymbol{x}}_S$, and $\boldsymbol{\sigma}_S$ of the forward motion watching at front direction.

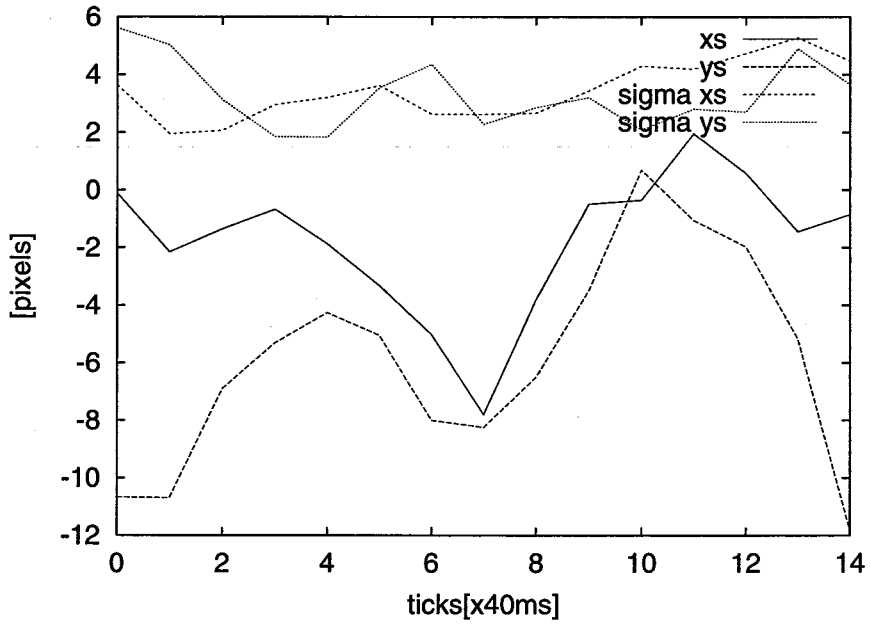


Fig.5.7 The $\hat{\boldsymbol{x}}_S$, and $\boldsymbol{\sigma}_S$ of the right forward motion watching at front direction.

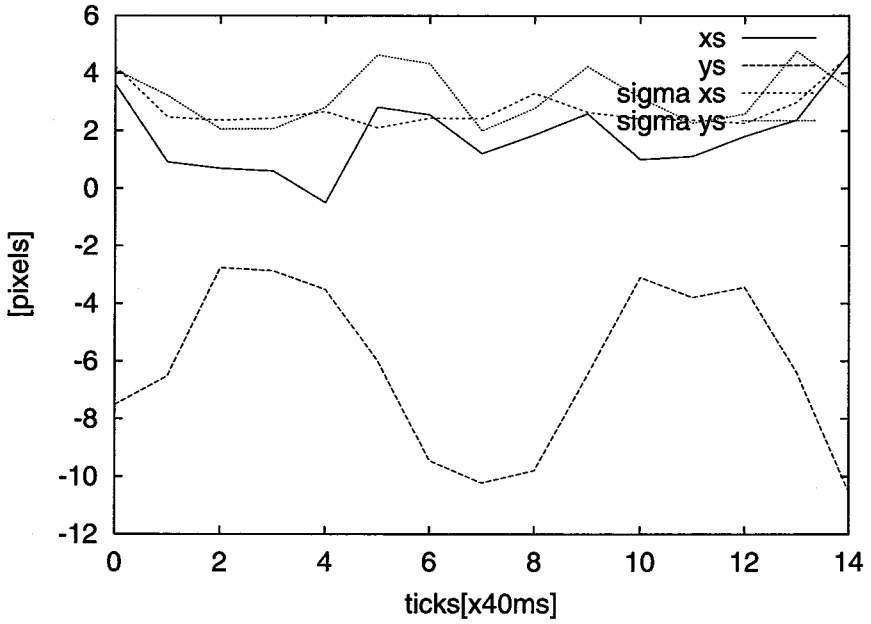


Fig.5.8 The $\hat{\boldsymbol{x}}_S$, and $\boldsymbol{\sigma}_S$ of the left forward motion watching at front direction.

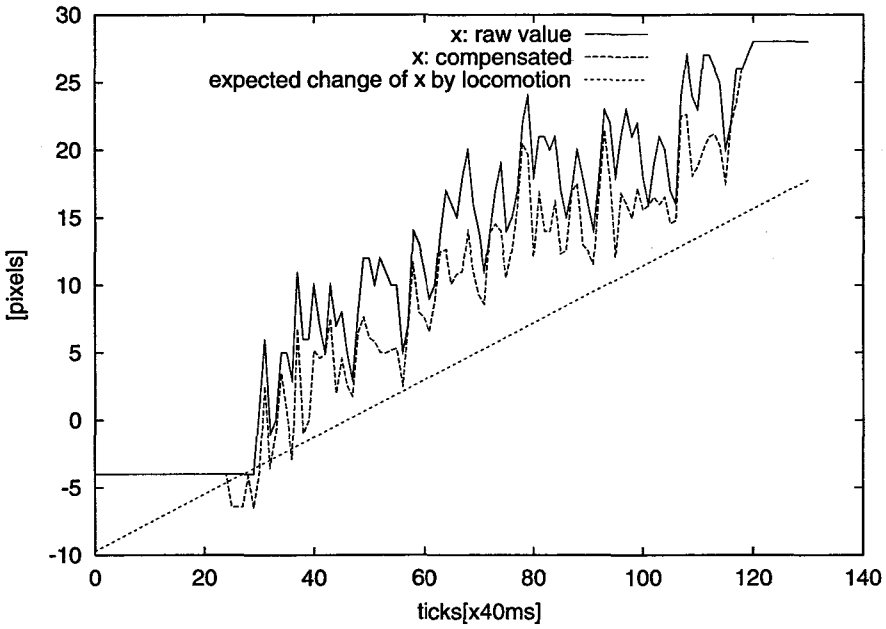


Fig.5.9 Image compensation to the x axis by proposed method.

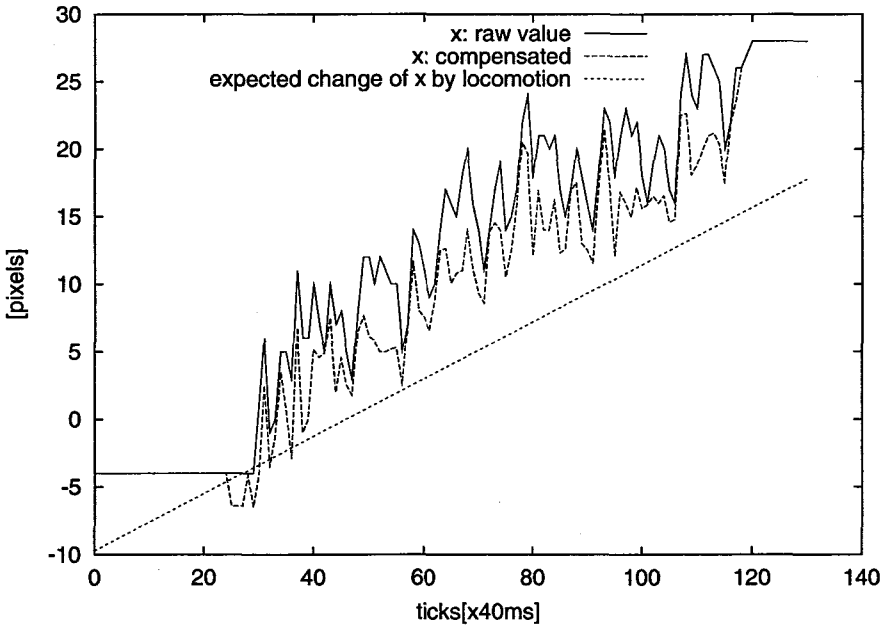


Fig.5.10 Image compensation to the y axis by proposed method.

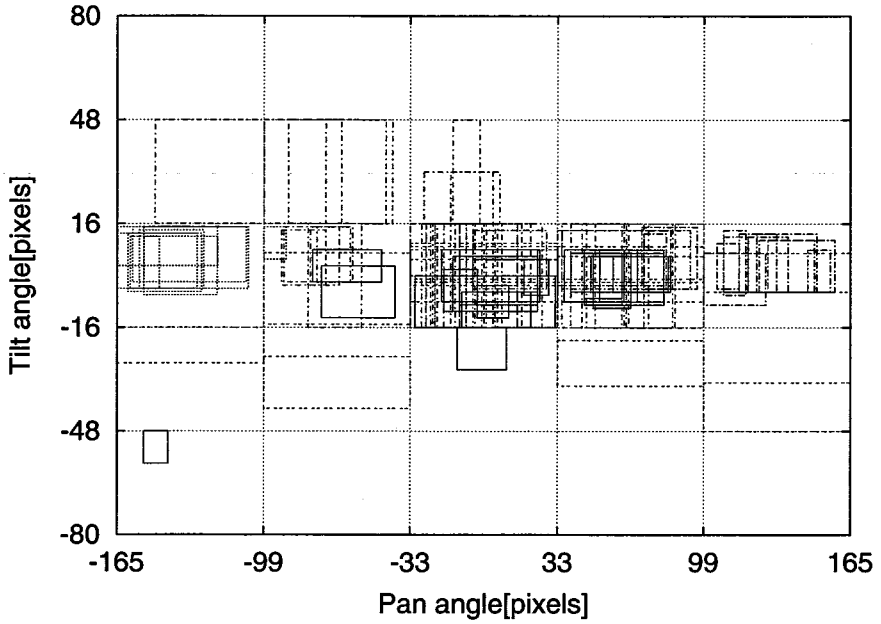


Fig.5.11 Generated attention windows by the proposed method.

5.4.3 提案手法による行動決定の様子

RoboCup 4足ロボットリーグ 2002 のフィールド (**Fig.** 5.3) において，ボールがゴール前にあるとき，ボールとゴールを一直線上に見る位置へ移動することをタスクとした．フィールド中央の3点のいずれかを初期位置として，トレーニングデータを集めた．フィールドが前の実験と比較して，広いため，ランドマークを観測するために，大きくパンする必要がある．フィールド中央の3点のいずれかを初期位置として，トレーニングデータを集めた．フィールド上には，進入禁止領域があり，ゴール正面から行動開始した場合には適切に避けなければならない．進入禁止領域は，Fig. 4.6 の中央の白い障害物を置いた領域である．ロボットは進入禁止領域についての視覚情報を持たず，ランドマーク等の観測から行動決定しなければならない．パン軸は 44 度毎の 5 方向に，チルト軸は 21 度毎の 5 方向に分けた．これはカメラの揺れに関わらず特定の方向を観測したときに，必ず行動決定木生成時に仮定した範囲が観測可能とするためである．616 のトレーニングデータを集め，情報量により行動決定木を生成した．**Fig.** 5.11 に生成された注視窓を示す．前の実験と比較して注視窓のある観測方向が広がっているが，正面に近い方向に注視窓が集中していることが分かる．観測前の行動に関するエントロピー (情報量期待値に等しい) は 1.8 となった．

Fig. 5.12 から **Fig.** 5.14 に，フィールド中央の点から行動開始したときの，観測による情報量の期待値，最大の行動確率とその行動の変化，静止観測した時期を示す．**Fig.** 5.12,

Fig. 5.13 では，情報量期待値の閾値は用いず，行動確率の閾値をそれぞれ 0.4 と，0.6 とした．**Fig.** 5.14 では，情報量期待値の閾値を 0.4，行動確率の閾値を 0.4 とした．

行動確率の閾値のみを用い閾値が 0.4 の場合 (**Fig.** 5.12) を見ると，行動確率が高くても，情報量期待値が 0.7 (観測前の情報量期待値の 4 割) 程度となることもあることが分かる．複数の試行を繰り返したところ，観測のために静止することが少なく，目標位置姿勢になるまでの時間も短かったが，フィールド上の進入禁止領域を通過した．このためタスクを達成したとは言えない．

進入禁止領域を正しく避けるよう行動確率の閾値を上げ，0.6 としたとき (**Fig.** 5.13) には，タスクは達成されたが，頻繁に観測のために静止し，目標位置姿勢になるまでの時間が長かった．7 試行し目標地点へ到達できたのは 6 試行であった．目標地点への移動時間は，6 試行の平均で 92[s]，観測のため静止していた時間は平均で 73%であった．

情報量期待値の閾値を用いた場合 (**Fig.** 5.14) には，行動確率の閾値は低いが，情報量期待値の閾値を満たさない場合には，静止観測する．移動行動に伴う行動確率の減少と情報量期待値の増大，観測に伴う行動確率の増大と情報量期待値の減少を繰り返しながら行動決定し，閾値を満たさないときには静止して観測している．歩行中の観測による情報量期待値の減少は，時刻 850 から 890 [×40ms] 付近に見られる．歩行中の情報量期待値の減少があまり見られないのは，観測方向の変更が移動による曖昧さの上昇に間に合わないためと考えられる．11 試行し目標地点へ到達できたのは 9 試行であった．目標地点への移動時間は，9 試行の平均で 52[s]，観測のため静止していた時間は平均で 55%であった．移動時間の平均は行動確率の閾値のみを高くした場合の 56%であり，情報量期待値の有効性を示している．

5.5 観測行動決定閾値についての考察

ここでは静止観測を行う条件として，2 つの指標，観測による情報量期待値と，行動確率の最大値のいずれかが満足されないときとしている．情報量期待値は，あるタスクを実行するための行動に関する，観測による情報量について求めている．よってタスクを実現するための行動についての曖昧さを表している．一方，観測確率から計算された行動確率には，観測の曖昧さは表現されていない．行動確率のエントロピーは，タスクを実現するための行動決定に関する曖昧さを表さない．

ここで実験中に見られた情報量期待値と行動確率の最大値の組を Fig.5.15 に示す．同じ情報量期待値であっても行動確率にバラツキが見られることがわかる．情報量期待値は閾値以下であるのに対し，行動確率が閾値以下となる場合も見られる．

情報量期待値が同じ値であっても，曖昧さを残しているノード (観測) により行動確率への影響が異なる原因の一つは，例えば木の上位の観測にある曖昧さを許す値が，木の下位では観測しない場合の値になる場合があるためである．行動決定木の下方にある観測ほどトレーニングデータ中に存在する割合が小さく，情報量期待値への寄与も小さいためで

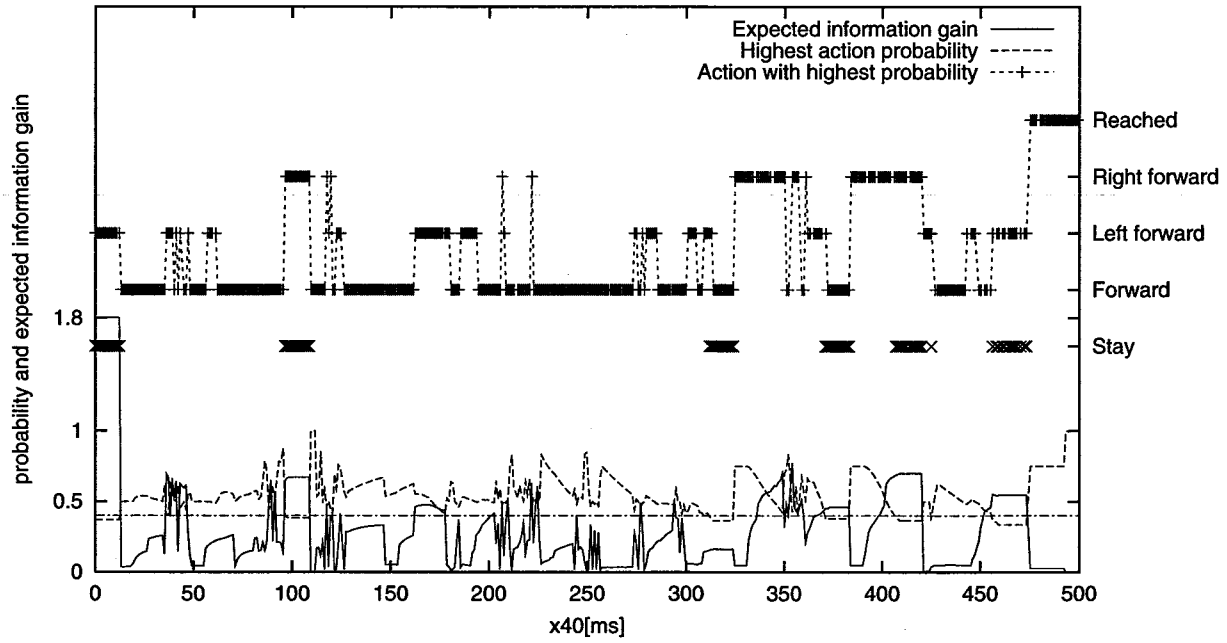


Fig.5.12 Changes of expected information gain, the highest action probability, and actions. One of which has the highest action probability, and the other is the one taken by the robot. The starting point of robot was the center of the field. The action probability threshold was 0.4 and the information gain threshold was not used.

ある．また観測を繰り返すことで，情報量期待値は単調減少するが，最大の行動確率は単調増加するとは限らない．これは行動確率の計算が非線形となっているためである．従って，Fig.5.15 のように情報量期待値が低い場合にも，行動確率が低いこともある．そこで本手法では最大の行動確率にも閾値を設けて，行動確率が低い場合には，静止観測を行っている．

静止観測のための指標としては他に，最大行動確率 p_{max} と現在の情報量期待値 I から計算した値，観測前の情報量を I_0 とし，最大行動確率に残る観測の曖昧さをかけた値が，閾値 p_0 を越えること，

$$p'_{max} = p_{max}(1 - \frac{I}{I_0})^n > p_0, \tag{5.8}$$

を指標として用いることも考えられる (n は任意の定数，Fig.5.15 に $n = 1, p_0 = 0.4, 0.6, 0.8$ のときの静止観測との境界を示す)．この指標を用いた実験を行ったが，前節での実験の結果と大きな差はなく，この平面上で，どのような指標を定義するべきであるかは今後の課題である．

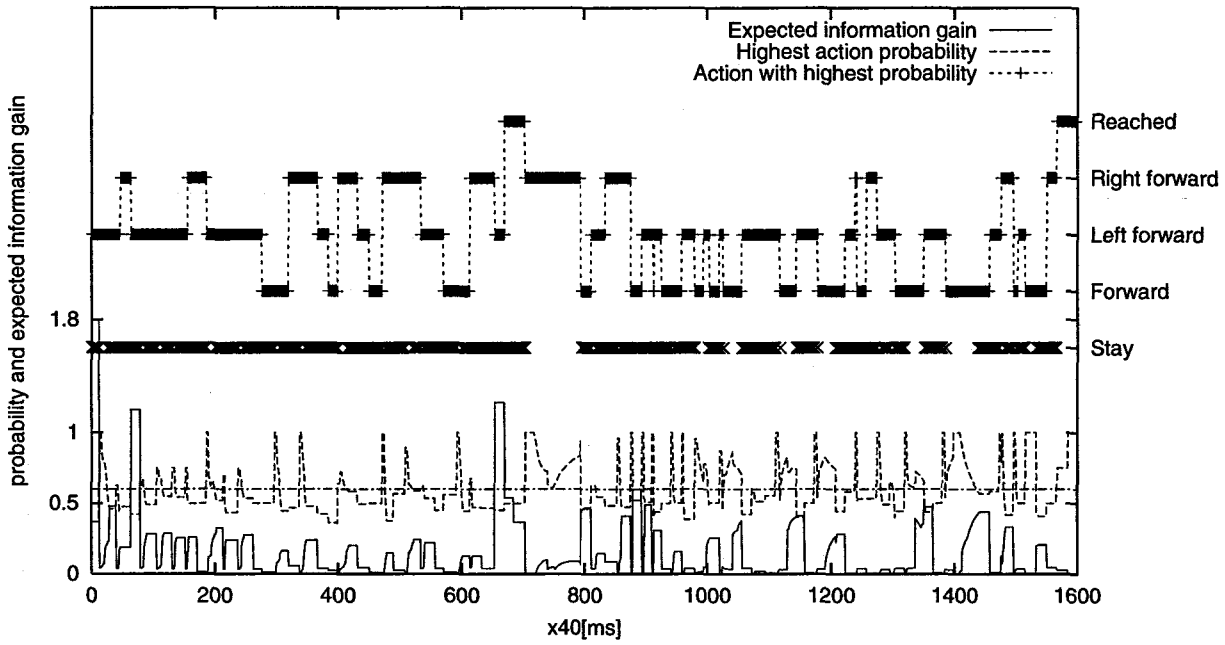


Fig.5.13 Changes of expected information gain, the highest action probability, and actions. One of which has the highest action probability, and the other is the one taken by the robot. The starting point of robot was the center of the field. The action probability threshold was 0.6 and the information gain threshold was not used.

5.6 結言

視覚センサを有する歩行ロボットのための，情報量に基づく注視戦略を提案し，観測による情報量期待値を用いることで行動決定の安定性を向上した．歩行による揺れを，視覚センサのみに基づいて補正し，移動中の観測を行動決定に用いることを実現した．実験に用いた歩行についての補正後の分散が比較的大きく，情報量期待値の増大が比較的早い実験結果となった．補正後の分散を低下するには，より高次の補正式を用いることや，運動学により歩行に伴うカメラの動きを補正すること，ジャイロなど他のセンサを併用すること，などが考えられる．本手法では歩行に応じた補正値を求める必要があるが，一方で歩行に関する知識を必要としないため，結果として滑りなどを含めた補正値となる利点がある．高速なフィードバック制御や，特別な機構を用いた揺れの補正法と比較すると，一般に分散は大きいと考えられるが，本手法を含む画像に基づく補正法は簡便で有用であると言える．

今後の課題としては，タスクにとって同じと見なせる複数の行動，たとえば遅いが揺れの少ない移動と，高速で揺れの激しい移動を，行動決定の効率によって自律的選択を行うことが挙げられる．

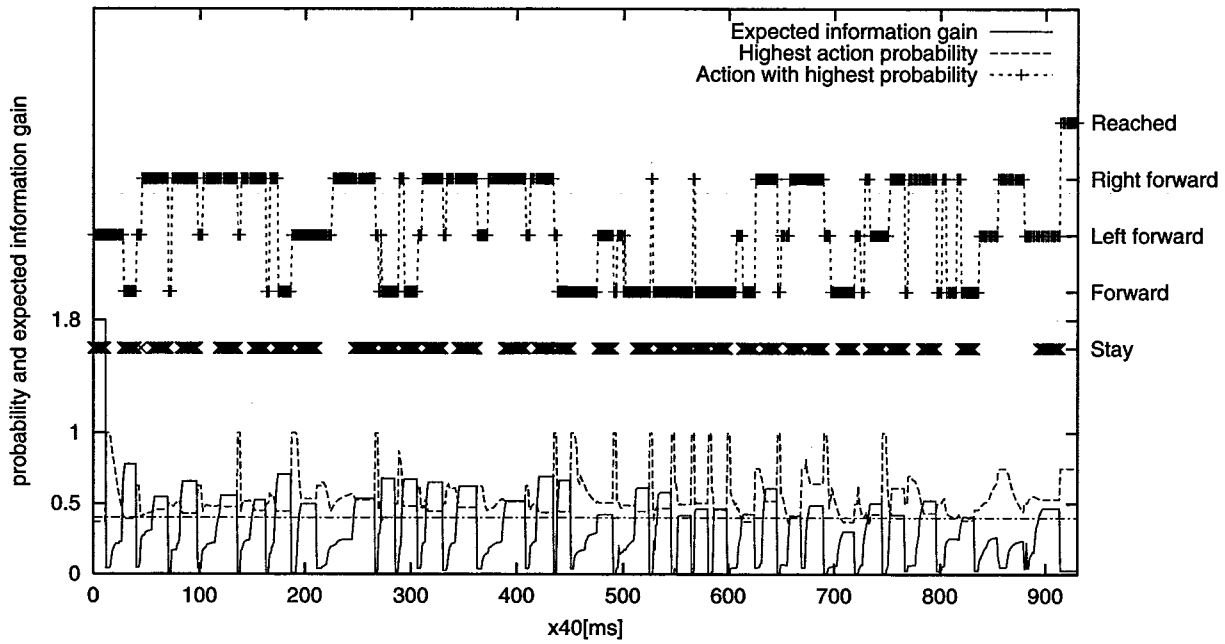


Fig.5.14 Changes of expected information gain, the highest action probability, and actions. One of which has the highest action probability, and the other is the one taken by the robot. The starting point of robot was the center of the field. The action probability threshold was 0.4 and the information gain threshold was 0.4.

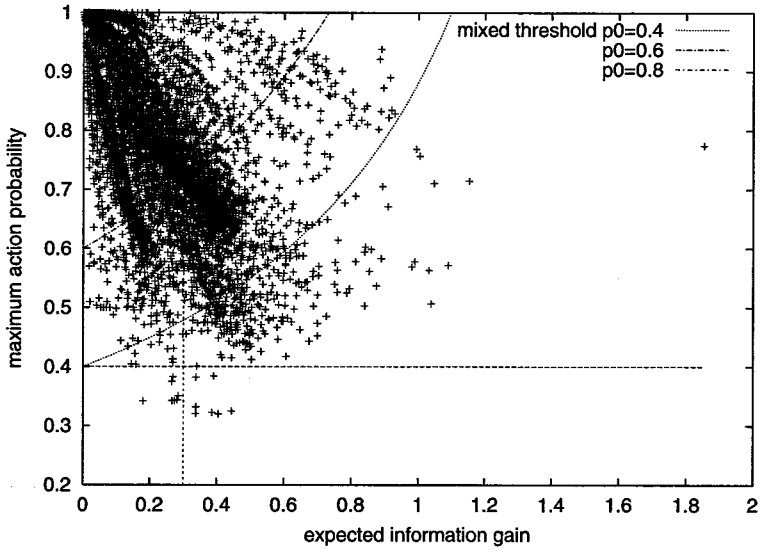


Fig.5.15 The values of expected information gain and maximum action probability observed in experiments.

第6章　結言

本論文では移動ロボットのための効率のよい観測を実現するための，情報量に基づく観測戦略を提案し，検証することを目的とした．

第1章，第2章では，移動ロボットと観測戦略に関する従来研究を概観し，それらにおける問題点を指摘し，本研究の目的を明確にした．すなわち，従来の移動ロボット研究でよく用いられる明示的な自己位置同定を行わない．自己位置同定のためや状態量推定のための観測戦略ではなく，タスクに応じた行動決定のための観測戦略を提案することである．

第3章では，情報量に基づく観測対象選択を行う観測戦略を提案した．観測をランドマークの方位を知ることとし，観測によるエントロピーの減少量，すなわち行動決定に関する観測の持つ情報量を教示データから計算する方法を示した．観測の持つ情報量は，観測による行動決定の曖昧さの減少量を示していることから，情報量の大きい観測から行うことで効率のよい観測が実現される．そこで行動に関する情報量に基づく行動決定木を生成し，行動決定のために教示データを圧縮し，同時に行動決定木の根から順に観測することで効率のよい観測が得られる．ここで観測対象は情報量により決定されるが，観測をランドマークの方位を知ることとしているため，観測方位は決定されない．そこで観測予測を行い，ランドマークの方位を予測すると共に，予測の曖昧なランドマークを観測する観測戦略を提案した．そして小型4脚ロボットを用いた実験により，手法の有効性を検証した．

第4章では，情報量によるセンサ空間の自律的分割と注視窓による観測対象と観測方向の選択を行う観測戦略を提案した．まずセンサ空間の自律的分割の重要性を実験データにより示した．つぎに，視覚センサの特性にあったセンサ空間分割法として，注視窓を提案した．観測をランドマークが注視窓内にあるか否かを知ることとし，ランドマークと注視窓の異なる観測候補から最も情報量の大きいものから順に観測することが，ここでの戦略である．これにより，第3章で提案した手法に対し，観測予測なしでの観測方向の決定，タスクに応じたセンサ値の離散化が可能となった．そして観測に必要な時間，例えば観測方向の変更に要する時間は観測により異なることから，情報量をその観測に必要な時間で割ることにより，単位時間辺りに得られる情報量 (情報量速度) を得る．情報量速度により行動決定のための決定木を生成し，決定木の根から順に観測する戦略を提案した．次に，一度生成した決定木の葉を状態とし，直前の状態が分かった場合に行動決定木を再生成することにより，行動決定に必要な観測を減らすことを提案した．そして実験により

提案手法の有効性を検証した．

第4章の手法を脚式移動ロボット適用し，移動中の観測を利用しより効率のよい行動決定をするためには，1) 歩行によるセンサの揺れの補正，2) 異なる観測位置でのセンサ値の統合，3) より適切な行動決定指標，が必要になる．そこで第5章では，歩行中の画像列から補正値を計算することを提案した．補正値の計算には画像列のみを用い，他のセンサは用いない．補正値を用いて，視覚センサによる観測から，静止時にある観測が得られる確率を計算することで，複数観測の統合と揺れの補正を実現した．また行動決定指標として，行動確率に加え，これから観測により得られると期待される情報量を行動決定の曖昧さの指標として用いることを提案した．行動確率あるいは情報量期待値が閾値を満足しない場合には，移動せずに観測する．提案した補正と情報量期待値により，脚式移動ロボットにおいて適切な行動決定が行われることを検証した．

今後の課題としては，教示負担の減少と動的環境への対応，教示データ追加時の部分的な行動決定木の生成，脚式ロボットの特徴をより生かした観測戦略などが挙げられる．教示負担の減少のためには，教示の自動化あるいは教示者に易しい形でのデータからロボットの視点への変換などが考えられる．教示データがあれば動的環境であっても同様に行動決定木の生成は可能であるが，1) 観測中に環境が変化してしまう，2) 教示データ収集のために周囲を観測している間に環境が変化してしまう，といった問題が考えられる．前者については，行動決定に情報が足りないため，センサを追加する，観測履歴を直前だけでなく，何時刻分か用いるといったことが考えられる．後者に関しては，教示中に情報の多い方向を観測する手法が望まれる．

教示データの追加に対する部分的な行動決定木の生成については，1) 生成された行動決定木の最適性，2) 部分生成に必要な教示データの圧縮記憶法，などの研究が望まれる．また，ここでは各状態において常に観測を繰り返すことで行動決定出来ると仮定したが，観測のために移動が必要な場合への拡張や，観測と移動のトレードオフの問題も課題である．

脚式ロボットに関する従来研究では，いかにして歩行を実現，獲得するかに重点がおかれ，目標移動速度の達成あるいは速い歩行の獲得が目指されてきた．即ち，脚の機能としては移動のみが考得られてきた．しかし，脚には移動だけでなくロボットの支持を行う機能がある．すなわち車輪式ロボットと違い，脚式ロボットでは姿勢を変えるために特別の機構を必要としない．ロボットの姿勢を変えることにより，観測視点をより広く変えることが可能となる．これによりカメラの角度を変えることでは同時に観測出来なかったランドマークを同時に観測できる可能性もあり，積極的に姿勢を変化させることで，より効率のよい観測が可能となる．また従来は歩行中の揺れは出来るだけ抑えることが目標とされてきたが，積極的に揺れを利用することにより，単眼カメラでのモーションステレオの利用も期待できる．

また観測だけでなくタスクにとっての効率性の視点も必要となる．観測に適した歩行を獲得，生成することも考えられるが，1) 観測に望ましい歩行が必ず実現できるとは限ら

ない，2) 脚先形状や，歩行面のちょっとした差などで影響される場合もあり，シミュレーションや軌道計画は容易でない，3) 実機による歩行の獲得が望ましいが効率のよい獲得法が確立されていない，といった問題を克服する必要がある．

付 録A　歩行の獲得

本章では，本論文の実験に用いた歩行の獲得法について述べる．ここでは，静歩行ではなく移動速度の高い動歩行の獲得を目指した．動歩行は主に4脚以下で実現される歩行で効率が良く，静歩行より移動速度が高いが，力の静的な釣合(静的安定性)が保たれず，ロボット固有のダイナミクスによる制御の難しい運動が生じる場合がある．その一つは，4脚ロボットの動歩行における，2脚期での支持直線回りの回転である．これまでは，動歩行を実現するために，1) この回転を極力押える, 2) 回転を無視したモデル化を行ない，実際との差はコンプライアンス制御により吸収する，という手法が主にとられてきた．前者の例として米田ら[71, 72]の間歇トロット歩容があげられる．これはZMPが支持直線上にのるようプランニングをすることで，2脚支持期の胴体の回転を極力抑えている．佐野ら[73]は後者の手法をとっており，回転を無視したモデルによりプランニングを行ない，モデル化誤差と着地時の衝撃をコンプライアンス制御により吸収する実験を行なっている．こういった手法を利用するためには，ある程度正確なモデル化とパラメータ，制御器とアクチュエータが必要となる．

それに対しアクチュエータを持たないロボットにおいて受動歩行により動歩行を実現する研究[74, 75, 76]が行なわれている．受動歩行ではロボット固有のダイナミクスによる運動を打ち消すように制御するのではなく，生かすことによって歩行を実現している．しかしアクチュエータを持たないため，重力などの外力がないと動作できず，歩行の柔軟性に欠ける．そこでアクチュエータを持つロボットにおいて，受動歩行と同様にロボットに固有のダイナミクスを利用した動歩行を実現することができれば，効率の高い歩行が比較的簡単な制御系により実現されることが期待される[77, 78]．またロボット自身の学習により，制御パラメータが獲得されることが望ましい．

以下，歩行獲得に関する従来研究を紹介し，4脚歩行ロボットを対象とした最急勾配法による歩行の獲得法を提案し，実ロボットによる獲得実験を報告する．また適切な歩行が実現されているかどうかの指標として安価な脚先接地センサを用いた指標である歩行実現度を提案する．

A.1 従来研究

Kunら[79]は，2脚のロボットでCMACを使った学習により動歩行を実現している．しかし，ダイナミクスを生かすのではなく，体を左右に振ることでダイナミクスの影響を減らすことによる実現となっている．

石黒ら[80]は，静歩行を行う4脚ロボットが，不整地での転倒防止のため脚の長さの調整法を遺伝的アルゴリズムで獲得する実験をシミュレーション上で行っている．橋本ら[81]は，車輪を持つ疑似4脚ロボットにおいて，疑似静歩行を実現するリズムジェネレータを遺伝的アルゴリズムによって獲得する実験を行っている．評価には移動速度と消費電流が含まれており，移動の効率も評価するようになっている．しかし，これらは静歩行を前提としており，ダイナミクスを有効に利用するとはいえない．

Gallagher and Beer[82]は，ナナフシのCPGを模した構造のニューラルネットワークのパラメータを遺伝的アルゴリズムで獲得する実験を，6脚ロボットのシミュレーションで行っている．田野ら[83, 84]は，確率的学習オートマトンにより，6脚ロボットの支持脚の上下動のタイミングの学習をシミュレーション上で行っている．Maes and Brooks[85]やTakita et. al[86]は，6脚の実ロボットにより，転倒しない反射的な移動の獲得実験を行っている．石黒ら[87]は，6脚の実ロボットにより，支持脚を遊脚へ切り替えるタイミングを決定する確率密度関数の学習実験をしている．これらの研究はいずれも，6脚以上での静歩行を何らかの形で仮定しており，ダイナミクスによる影響を避けている．

近藤ら[88]，Hornbyら[89]は4脚ロボットの実機による歩行の獲得実験をしている．いずれの場合も歩行面に強く依存した歩行が獲得されており，環境とロボットのダイナミクスを生かした歩行と考えられる．近藤ら[88]は，脚先位置をいくつか用意し，それらの組合わせと時間的な切り替えを遺伝的アルゴリズムにより獲得し歩行を実現した．Hornbyら[89]は，歩行に必要であると思われる脚先軌道を生成できる関数を用意し，その関数のパラメータを遺伝的アルゴリズムにより獲得した．前者に対し，後者はパラメータが多く，その学習時間も長くかかったことが報告されている．

A.2 ダイナミクスを生かした歩行獲得

本研究では環境とロボットのダイナミクスを有効に利用した，4脚歩行ロボットの動歩行の獲得法を提案する．全てのダイナミクスをシミュレーション上に記述することは困難であるため，実ロボットによる獲得ができるものとする．そこで獲得に必要なパラメータを最小限にし，実験の負担を軽くすることを考える．またフィードバック制御は行わず，フィードフォワード制御のみを用いる．

A.2.1 運動モデルと歩行パラメータ

4脚での動歩行は，対角脚が組となるトロット，進行方向に対して前後となる脚が組となるペース，進行方向に対して左右となる脚が組となるバウンスに分けられる．いずれも2脚支持を基本としており，2脚期に正確な制御を行うには大きな足首トルクと足裏摩擦が必要となる．以下では足首にアクチュエータを持たないロボットにおける，定常な歩行での2脚期の運動を簡単なモデルで表し，獲得すべき歩行パラメータについて考察する．トロットを仮定するが，ペース，バウンスについても同様の議論が可能である．

2脚期の4脚ロボットを，倒立振子で近似する．支持脚間を結んだ直線と支持直線と重心の距離を l とおく．ロボットの脚は十分軽く質量は胴体に集中し，歩行中脚のつけねから脚先までの長さがさほど変化しないと仮定すると，l は一定とみなせる．このときの重心の動きは倒立振子となるので，遊脚を能動的に降ろさなかった際に着地するまでにかかる時間は，重心の位置 l と重心の初期姿勢 θ_1，初期速度 $\dot{\theta_1}$，遊脚の脚先高さ $\triangle h$ に依存し，

$$t_L = t_L(l, \triangle h, \theta_1, \dot{\theta_1}) \tag{A.1}$$

と書ける．$\triangle h$ が大きいほど，$\theta_1, \dot{theta_1}$ が小さいほど t_L は長くなる．遊脚を能動的に降ろす際の遊脚着地までの時間 t_l は，$t_l < t_L$ であればよいが，θ_1 が同じであれば重力により加速され t_l が長いほど，遊脚着地時の重心の速度は向上する．一方 t_l が t_L を越えると，遊脚期の脚軌道を描いている脚が着地することになり歩行速度が低下する．したがって，移動速度は t_l に対して t_L を山とする単峰性の曲線を描く．逆に t_L に対し移動速度は t_l を山とする単峰性の曲線を描くと考えられる．

初期姿勢 θ_1 は新しい2脚支持脚期の開始時，すなわち遊脚着地点で決まる．遊脚開始点を結んだ直線(支持対角線)を，胴体上に固定した直線に対する回転角 α と，距離 x で表す．重心のこの直線に対する距離を d とすると，

$$(d - x)\cos\alpha = l\sin\theta_1 \tag{A.2}$$

$$\theta_1 = \arcsin\frac{(d - x)\cos\alpha}{l} \tag{A.3}$$

α, x は θ_1, t_L と単調な関係にあることがわかる．t_l を固定し x, α を変化させた場合，移動速度は単峰性を示すと考えられる．

α は支持対角線の進行方向に対する角度を決定し，小さいほど進行方向と倒立振子の倒れる方向が近くなる．α が大きいと速度が低下し，限界を越えるとロボットは左右に胴体を振るのみとなり前進しない．ある範囲内では α は小さい方が移動速度は向上するが，小さ過ぎると遊脚着地時の前後方向の衝撃が大きくなり移動速度は低下し，さらには転倒すると考えられる．

A.2.2 歩行パラメータの探索法

遺伝的アルゴリズムのような大域探索ではなく，$l, \triangle h$ は固定した，次のような最急勾配法による局所探索とした．

1. t_l, α を実現可能と思われる値に固定し，最急勾配法で移動速度を最大にする x を求める
2. t_l, x を固定し，最急勾配法で移動速度を最大にする α を求める
3. t_l を 増加させ，1 から繰り返す, 移動速度が向上しない場合には終了

A.2.3 脚先接地センサによる歩行実現度

ロボット自身が移動速度を知るためには，1) 距離センサ, 2) カメラ画像, 3) 加速度センサなどを使う方法が考えられる．距離センサやカメラ画像を使う場合には環境の整備と，センサのキャリブレーションなどが必要となる．加速度センサを利用する場合には，環境の整備は不要であるが，ノイズが多く，一定速度で移動する場合には使えない．

それに対し，脚先の滑べりが無視できる場合に，脚先接地センサを使った歩行実現度を提案する．脚先が適切な軌道を描いており，支持脚，遊脚が適切に切り替われば歩行は実現される．切替が不適切であれば，遊脚であるべき脚が着地し，同時に支持脚であるべき脚が遊脚となる．そこで一定の期間歩行した際の脚先接地センサの値を利用した値，

$$(\text{歩行実現度}) = \frac{(\text{測定期間}) - (\text{1 脚でも遊脚が着地した期間})}{(\text{測定期間})} \tag{A.4}$$

を提案する．この値が高いほど脚先軌道に無理がなく，支持脚，遊脚の切替が滑らかに行なわれ，移動速度が高いと期待される．

移動速度そのものが測定されるわけではなく，脚先の滑べりの影響や，受動歩行の重力による加速の効果は測れない ($t_l > t_L$ による速度低下は測定されるが，$t_l < t_L$ における速度向上を測定されない)．しかし，歩行実現度は，安価なロボットの内界センサのみを用いた歩行の評価基準となり，前進だけでなく，回転，左右移動など歩行一般に用いることができる．また歩行実現度をロボットの左右独立に計算することで，ロボットの直進性も評価できると期待される．

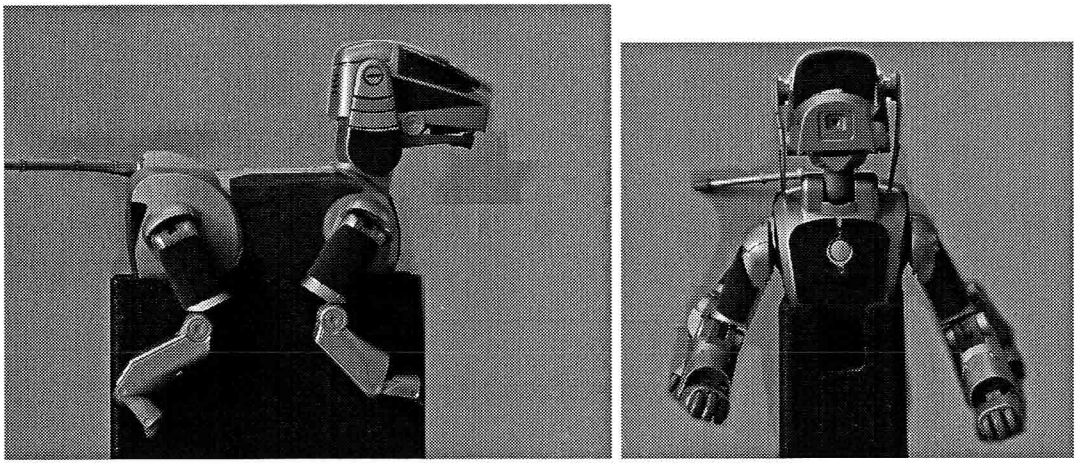
A.3 実験結果

A.3.1 ロボットの構造と脚先軌道

実験に用いたロボットを横と正面から見たところを Fig. A.1 に示す．各脚に 3 自由度を持つ 4 脚ロボットで，脚の付け根 (第 1 軸) と膝に相当する第 3 軸は前後に回転し，第 2 軸

はそれらに垂直な軸周りに回転する．脚先には接地を検出するスイッチとバネの入った爪先がある．

各関節は位置制御されているが，自重により目標値を必ずしも達成できるとは限らない．脚先軌道は，軌道中の4点(脚先が最も前となる点，下となる点，後となる点と上となる点)のみを指定し，4点間の軌道については関節の位置制御にまかせることとした．



(a) side view (b) front view

Fig.A.1 The side and front view of the legged robot

A.3.2 歩行パラメータの変動と歩行実現度, 移動速度

脚先軌道と脚の前後間隔，歩行周期を一定とし，前後の脚の遊脚着地点を同量だけ前から後ろに移動した．このときは θ_1 のみが変化すると考えられる．遊脚着地点の位置に対して，実現された移動速度と歩行実現度を Fig. A.2 に示す．

次に脚先軌道と歩行周期を一定とし，前脚と後脚の遊脚着地点を同量ずつ逆に移動し，支持対角線の進行方向に対する傾きを変えた．このとき α のみ変化し，θ_1 は変化しないと考えられる．支持対角線の傾きに対して，実現された移動速度と歩行実現度を Fig. A.3 に示す．前後脚の間隔が広く移動速度が遅い，支持直線の傾きが 0.9[rad] 以上の領域では数ステップで左右に胴体を振動するモードに入り前進しなくなってしまう挙動を示した．逆に前後脚の間隔を狭くしていくと，支持直線の傾きが 0.73[rad] 以下の領域では安定性に欠け前に倒れ込む挙動を示した．また脚間隔が狭い条件で支持直線の位置によっては後に倒れ込む挙動も見られた．

同様に，歩行周期を変化させた場合も単峰性が見られた．これらより，局所探索に最急

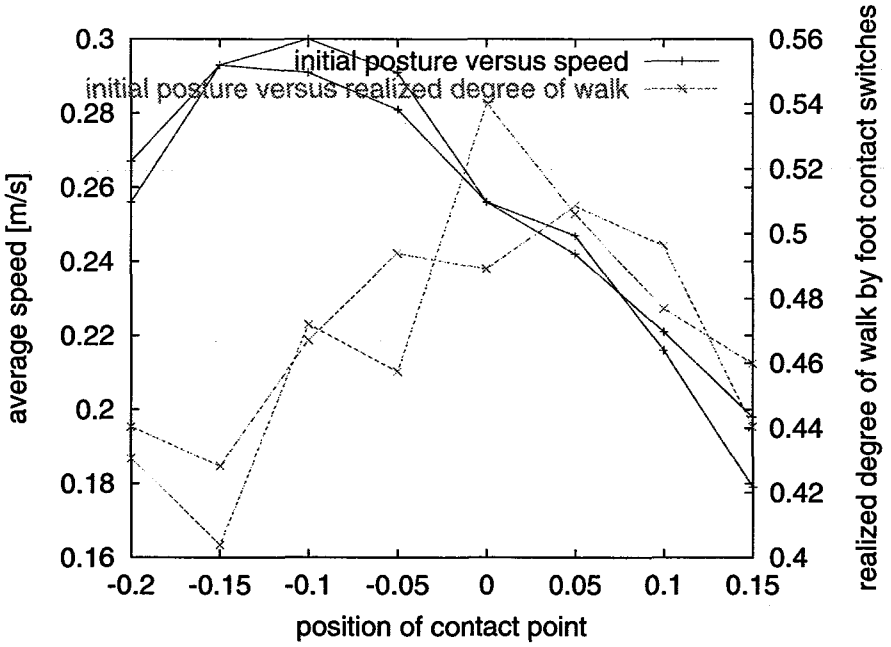


Fig.A.2 Landing point of leg versus walk speed and realized degree of walk.

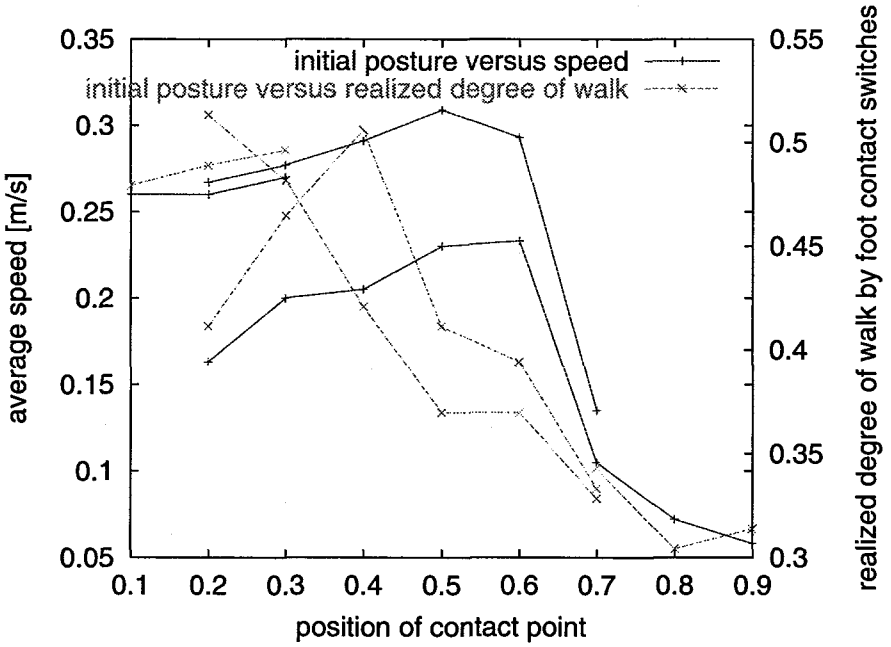


Fig.A.3 Angle of supporting line versus walk speed and realized degree of walk.

勾配法を用いることが有効であることが分かったので，脚先軌道をいくつか用意し，第3章，第4章で用いた歩行を最急勾配法により探索し求めた．最終的な歩行を得る過程で，カーペットの織目に沿った方向で歩行を獲得すると，逆方向や垂直方向ではほんとんど移動しない，ロボットと歩行面に強く依存した歩行となる現象も見られた．

A.3.3 歩行実現度と移動速度

実験により得られた歩行実現度と移動速度の関係をFig.A.4に示す．この図から正の相関があることがわかる．歩行実現度が1にならず，歩行実現度と移動速度の山が一致しない(Fig. A.2, Fig. A.3)理由の一つは，脚先接地センサの構造上の問題があげられる．脚先接地センサが反応するにはある程度の押し込みが必要であるが，この押し込みの範囲では脚先圧力は低く歩行を妨げないのに接地しているとされている可能性がある．また実験に用いたロボットには爪があり，脚先接地センサより先に着地する可能性がある．しかし，移動速度との正の相関が認められある程度の指標として使えること考えられる．どの程度役に立つ指標であるかは今後の研究課題である．

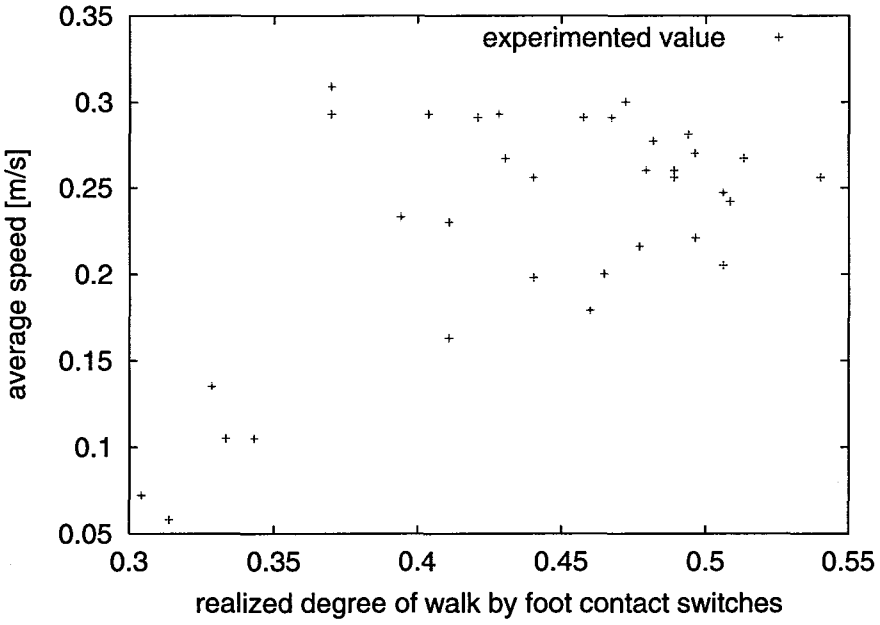


Fig.A.4 Realized degree of walk by foot contact switches versus average speed

参考文献

[1] 石黒浩. 注視に基づくロボットの視覚. 日本人工知能学会誌, Vol. 10, No. 4, pp. 500–506, 1995.

[2] Alberto Elfes. Sonar-based real-world mapping and navigation. *IEEE Journal of Robotics and automation*, Vol. RA-3, No. 3, pp. 249–265, 1987.

[3] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. A probabilistic approach to concurrent mapping and localization for mobile robots. *Machine Learning*, Vol. 31, pp. 29–53, 1998.

[4] 辻安彦, 八木康史, 谷内田正彦. 全方位視覚センサを用いたロバストな環境マップ生成と自己位置推定. 日本ロボット学会誌, Vol. 19, No. 1, pp. 59–67, 2001.

[5] Shoichi MAEYAMA, Akihisa OHYA, and Shin'ichi YUTA. Outdoor navigation of a mobile robot using natural landmarks. In Y. Kakazu, M. Wada, and T. Sato, editors, *In Proc. of the Intelligent Autonomous Systems 5*, pp. 164–171, 1998.

[6] Tom Duckett and Ulrich Nehmzow. Exploration of unknown environments using a compass, topological map and neural network. In *In Proc. CIRA '99, 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation*, 1999.

[7] Ulrich Nehmzow, Toshihiro Matsui, and Hideki Asoh. "virtual coordinates": Perception-based localisation and spatial reasoning in mobile robots. In Y. Kakazu, M. Wada, and T. Sato, editors, *In Proc. of the Intelligent Autonomous Systems 5*, pp. 188–195, 1998.

[8] 藤澤加絵, 早川聡一郎, 青木猛, 鈴木達也, 大熊繁. 自律移動ロボットにおける実時間行動探索. 日本ロボット学会誌, Vol. 17, No. 4, pp. 503–512, 1999.

[9] 吉岡孝, 登尾啓史, 富永昌治. 移動ロボットのセンサベーストナビゲーションについて. 日本ロボット学会誌, Vol. 19, No. 8, pp. 991–1002, 2001.

[10] S. D. Whitehead. A complexity analysis of cooperative mechanisms in reinforcement learning. In *Proceedings of AAAI-91*, pp. 607–613, 1991.

[11] Jun Tani. Model-based learning for mobile robot navigation from the dynamical systems perspective. *IEEE Trans. on System, Man and Cybernetics Part B (Special Issue on Robot Learning)*, Vol. 26, No. 3, pp. 421–436, 1996.

[12] Jun Tani. A dynamical systems approach to represent cognition of robots: A view of the internal observer. In *AAAI Fall Symposium, Embodied Cognition and Action*, pp. 123–128, 1996.

[13] Yuko ISHIWAKA and Masaaki MINAGAWA. Training mobile agent to solve cellular maze problem: An ann based approach. In Y. Kakazu, M. Wada, and T. Sato, editors, *In Proc. of the Intelligent Autonomous Systems 5*, pp. 180–187, 1998.

[14] 鈴木育男, 横井浩史, 嘉数侑昇. 連想記憶を用いたロボットナビゲーションシステムに関する基礎研究. 第17回日本ロボット学会学術講演会講演論文集, pp. 613–614. 日本ロボット学会, 1999.

[15] Jun Tani, Jun Yamamoto, and Hiro Nishi. Dynamical interactions between learning, visual attention, and behavior: An experiment with a vision-based mobile robot. In Phil Husbands and Inman Harvey, editors, *Fourth European Conference on Artificial Life*, pp. 309–317. The MIT Press, 1997.

[16] N. Ayache and O.D. Faugeras. Maintaining representations of the envrironment of a mobile robot. *IEEE Transactions on Robotics and automation*, Vol. 5, No. 6, pp. 804–819, 1989.

[17] D. J. Kriegman, E. Triendl, and T. O. Binford. Stereo vision and navigation in buildings for mobile robots. *IEEE Transactions on Robotics and automation*, Vol. 5, No. 6, pp. 792–803, 1989.

[18] Tom Duckett and Ulrich Nehmzow. Mobile robot self-localisation and measurement of performance in middle-scale environments. *Robotics and Autonomous Systems*, Vol. 1-2, No. 24, pp. 57–69, 1998.

[19] Tom Duckett and Ulrich Nehmzow. Knowing your place in real world environments. In *Proc. EUROBOT '99, 3rd European Workshop on Advanced Mobile Robots*. IEEE Computer Society, 1999.

[20] J. A. Castellanos, J. M. M. Montiel, J. Neira, and J. D. Tardòs. Sensor influence in the performance of simultaneous mobile robot localization and map building. In *Preprints for 6th International Symposium on Experimental Robotics*, pp. 203–212, 1999.

[21] M. W. M. G. Dissanayake, P. Newman, H. F. Durrant-Whyte, S. Clark, and M. Csorba. An experimental and theoretical investigation into simultaneous localisation and map building. In *Preprints for 6th International Symposium on Experimental Robotics*, pp. 171–180, 1999.

[22] John J. Leonard and Hans Jacob S. Feder. Experimental analysis of adaptive concurrent mapping and localization using sonar. In *Preprints for 6th International Symposium on Experimental Robotics*, pp. 213–222, 1999.

[23] Patrick Rives, José-Luís Sequeira, and Pedro Lourtie. Continuous probabilistic mapping. In *Preprints for 6th International Symposium on Experimental Robotics*, pp. 181–190, 1999.

[24] Akio Kosaka and Avi Kak. Fast vision-guided mobile robot navigation using model-based reasoning and prediction of uncertinties. In *Proceedings of the 1992 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 3, pp. 2177–2186, 1992.

[25] Akio Kosaka, Min Meng, and A. C. Kak. Vision-guided mobile robot navigation using retroactive updating of position uncertainty. In *Proceedings of the 1993 IEEE International Conference on Robotics and Automation*, Vol. 2, pp. 1–7, 1993.

[26] 前山祥一, 大矢晃久, 油田信一. 移動ロボットのための遡及的現在位置推定法-処理時間を要する外界センサデータの利用. 日本ロボット学会誌, Vol. 15, No. 7, pp. 1075–1081, 1997.

[27] Anthony R. Cassandra, Leslie Pack Kaelbling, and Michael L Littman. Acting optimally in partially observable stochastic domains. In *Proc. of the twelfth International Conference on Artificial Intelligence*, pp. 1023–1028, Seattle, WA, 1994.

[28] Reid Simmons and Sven Koenig. Probabilistic robot navigation in partially observable environments. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1080–1087, 1995.

[29] Anthony R. Cassandra, Leslie Pack Kaelbling, and James A. Kurien. Acting under uncertainty: Discrete bayesian models for mobile robot navigation. In *Proceedings*

of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 963–972, 1996.

[30] W. Burgard, D. Fox, D. Hennig, and T. Schmidt. Estimating the absolute position of a mobile robot using position probabilitygrids. In *Proc. of the National Conferenceon Artificial Intelligence*, pp. 896–901, 1996.

[31] Wolfram Burgard, Andreas Derr, Dieter Fox, and Armin B. Cremers. Integrating global position estimation and position tracking for mobile robots: The dynamic mrkov localization approach. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 730–735, 1998.

[32] Jens-Steffen Gutmann, Wolfram Burgard, Dieter Fox, and Kurt Konolige. An experimental comparison of localization methods. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 736–743, 1998.

[33] S. Thrun, D. Fox, W. Burgard, and Dellaert. F. Robust monte carlo localization for mobile robots. *Artificial Intelligence*, Vol. 128, No. 1-2, 2001.

[34] Stefan Enderle, Marcus Ritter, Dieter Fox, Stefan Sablatnög, Gerhard Kraetzschmar, and Günther Palm. Vision-base localization in robocup environments. In *Proc. of The Fourth International Workshop on RoboCup*, pp. 232–237, 2000.

[35] Scott Lenser and Manuela Veloso. Sensor resetting localization for poorly modeled mobile robots. In *Proceedings of ICRA-2000*, 2000.

[36] 上田隆一, 小林祐一, 深瀬武, 新井民夫, 湯浅秀男, 太田順. 一様分布に基づく高速モンテカルロ自己位置同定. 第 19 回日本ロボット学会学術講演会講演予稿集, pp. 1007–1008. 日本ロボット学会, 2001.

[37] Andrew Kachites McCallum. Learning to use selective attention and short-term memory in sequential tasks. In *From animals to animates 4*, pp. 315–324. 1995.

[38] Yasutake Takahashi, Minoru Asada, and Koh Hosoda. Reasonable performance in less learning time by real robot based on incremental state space segmentation. In *Proceedings of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1518–1524, 1996.

[39] Takehisa Yairi, Shinichi Nakasuka, and Koichi Hori. State abstraction from heterogeneous and redundant sensor information. In Y. Kakazu, M. Wada, and T. Sato, editors, *In Proc. of the Intelligent Autonomous Systems 5*, pp. 234–241, 1998.

[40] Masatoshi KAMIHARAKO, Hiroshi ISHIGURO, and Toru ISHIDA. Attention control for state space construction. In Y. Kakazu, M. Wada, and T. Sato, editors, *In Proc. of the Intelligent Autonomous Systems 5*, pp. 258–265, 1998.

[41] Minoru Asada, Shoichi Noda, and Koh Hosoda. Action based sensor space segmentation for soccer robot learning. *Applied Artificial Intelligence*, Vol. 12, No. 2-3, pp. 149–164, 1998.

[42] 矢入健久, 中須賀真一, 堀浩一. 異種冗長なセンサー情報に基づく自律的状態抽象化法. 日本人工知能学会誌, Vol. 14, No. 4, pp. 667–678, 1999.

[43] Takehisa Yairi, Koichi Hori, and Shinichi Nakasuka. Autonomous reconstruction of state space for learning of robot behavior. In *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 891–896, 2000.

[44] 矢入健久, 堀浩一, 中須賀真一. 複数行動結果を考慮した最尤推定に基づく状態一般化法. 日本人工知能学会誌, Vol. 16, No. 1, pp. 128–138, 2001.

[45] 文仁赫, 三浦純, 白井良明. 不確かさを考慮した観測位置と移動のオンライン計画手法. 日本ロボット学会誌, Vol. 17, No. 8, pp. 1107–1113, 1999.

[46] Kiyosumi Kidono, Jun Miura, and Yoshiaki Shirai. Autonomous visual navigation of a mobile robot using a human-guided experience. In *In Proc. of the Intelligent Autonomous Systems 6*, pp. 620–627. IOS Press, 2000.

[47] Kazunori Takeuchi, Jun Ota, Kazuo Ikeda, Yasumichi Aiyama, and Tamio Arai. Mobile robot navigation using artificial landamarks. In *In Proc. of the Intelligent Autonomous Systems 6*, pp. 431–438. IOS Press, 2000.

[48] 城殿 清澄, 三浦純, 白井良明. 誘導による移動経験に基づく視覚移動ロボットの自律走行. 日本ロボット学会誌, Vol. 19, No. 8, pp. 1003–1009, 2001.

[49] 三浦純, 白井良明. プランニングコストと視覚の不確かさを考慮した移動ロボットの視覚と行動のプランニング. 日本人工知能学会誌, Vol. 13, No. 4, pp. 588–596, 1998.

[50] Patric Jensfelt, David Austin, and Henrik I. Christensen. Towards task oriented localization. In *In Proc. of the Intelligent Autonomous Systems 6*, pp. 612–619. IOS Press, 2000.

[51] R. A. Howard. ダイナミックプログラミングとマルコフ過程. 培風館, 1971.

[52] Hotaka Takizawa, Yoshiaki Shirai, Jun Miura, and Yoshinori Kuno. Plannning of observation and motion for interpretation of road intersection scenes considering uncertainty. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 520–525, 1998.

[53] Anthony R. Cassandra, Michael L. Littman, and Nevin L. Zhang. Incremental pruning: A simple, fast, exact method for partially observable markov decision processes. In *Uncertainty in Artificial Intelligence (UAI)*. 1997.

[54] Michael Littman, Anthony Cassandra, and Leslie Kaelbling. Learning policies for partially observable environments: Scaling up. In *Proceedings of the Twelfth International Conference on Machine Learning*, pp. 362–370. Morgan Kaufmann, 1995.

[55] Kazuteru MIYAZAKI and Shigenobu KOBAYASHI. Learning deterministic policies in partially observable markov decision processes. In Y. Kakazu, M. Wada, and T. Sato, editors, *In Proc. of the Intelligent Autonomous Systems 5*, pp. 250–257, 1998.

[56] Michael Kearns, Yishay Mansour, and Andrew Y. Ng. Aproximate plannning in large pomdps via reusable trajectories. 1999.

[57] Steven D. Whitehead and Dana H. Ballard. Active perception and reinforcement learning. In *Proceedings of the Seventh International Conference on Machine Learning*, pp. 179–188. Morgan Kaufmann, 1990.

[58] Andrew J. Davison and David W. Murray. Mobile robot localisation using active vision. In *Proc. 5th European Conf. on Computer Vision*, pp. 809–825, 1998.

[59] 鈴木孝幸, 横井浩史, 嘉数侑昇. 情報重要度に基づくセンシングの効率化に関する研究. 第18回日本ロボット学会学術講演会講演予稿集, pp. 1001–1002. 日本ロボット学会, 2000.

[60] Seth A. Hutchinson and Avinash C. Kak. Planning sensing strategies in a robot work cell with multi-sensor capabilities. *IEEE Transactions on Robotics and automation*, Vol. 5, No. 6, pp. 765–783, 1989.

[61] Yutaka Sakaguchi. Haptic sensing system with active perception. *Advanced Robotics*, Vol. 8, No. 3, pp. 263–283, 1994.

[62] 阪口豊. 動きの予測を伴う能動的認識のアルゴリズム. 日本ロボット学会誌, Vol. 12, No. 4, pp. 708–714, 1994.

[63] W. Burgard, D. Fox, and S. Thrun. Active mobile robot localization. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI).* Morgan Kaufmann, San Mateo, CA, 1997.

[64] Takeshi Fukase, Yuichi Kobayashi, Ryuichi Ueda, and Tamio Arai. Real-time decision making under uncertainty of self-localization results. 人工知能学会 第15回 SIG-Challenge 研究会, pp. 21–26. 2002.

[65] 宮下敬宏, 細田耕, 竹内進, 浅田稔. 反射の組合わせによる歩行の実現. 日本ロボット学会誌, Vol. 18, No. 3, pp. 381–386, 2000.

[66] 岡田慧, 加賀美聡, 稲葉雅幸, 井上博允. 4脚歩行ロボット JROB-1 の視覚に基づく行動制御. ロボティクス・メカトロニクス'99 講演会講演論文集, pp. 1A1–63–096. 日本機械学会, 1999.

[67] 原光博, 福田正幹, 西林ひなた, 相山康道, 太田順, 新井民夫. 歩行時の揺れモデルを持つ4脚ロボットによる協調搬送制御. ロボティクス・メカトロニクス'99 講演会講演論文集, pp. 2P2–56–086. 日本機械学会, 1999.

[68] J. Ross Quinlan. Discovering rules from large collections of examples: a case study. In D. Michie, editor, *Expert Systems in the Microelectronic Age.* University Press, Edinburgh, Scotland, 1979.

[69] Manuela Veloso, Hiroaki Kitano, Enrico Pagello, Gerhard Kraetzshmar, Peter Stone, Tucker Balch, Minoru Asada, Silvia Coradeschi, Lars Karlsson, and Masahiro Fujita. Overview of robocup-99. In Manuela Veloso, Enrico Pagello, and Hiroaki Kitano, editors, *RoboCup-99: Robot Soccer World Cup III*, pp. 1–34. Springer, Lecture Note in Artificail Intelligence (1856), 2000.

[70] J. Ross Quinlan. *C4.5: PROGRAMS FOR MACHINE LEARNING.* Morgan Kaufmann Publishers, 1993.

[71] Kan YONEDA, Hiroyuki IIYAMA, and Shigeo HIROSE. Intermittent trot gait of a quadruped walking machine dynamic stability control of and omnidirectional walk. In *Proceedings of the 1996 IEEE International Conference on Robotics and Automation*, pp. 3002–3007. IEEE, 1996.

[72] 米田完, 飯山浩幸, 広瀬茂男. 4足歩行機械の全方向トロット歩容. 第12回ロボット学会学術講演集, pp. 1229–1230, 1994.

[73] 佐野明人, 古荘純次, 橋口昭浩. コンプライアンス制御機能を有する 4 足歩行システムに関する基礎的研究. 日本機械学会論文集 C 編, Vol. 57, No. 539, pp. 2297–2304, 1991.

[74] 中西秦人, 浜田陽一, 下山勲, 三浦宏文. 四足受動歩行ロボット. 第 6 回知能移動ロボットシンポジウム, pp. 21–26, 1992.

[75] 田口幹, 金澤正樹, 釣谷路人, 阿久根圭, 高津英生. 疑似直線機構を用いた 4 足受動歩行機械. 第 10 回日本ロボット学会学術講演会講演論文集, pp. 567–568. 日本ロボット学会, 1992.

[76] Adam C. Smith and Matthew D. Berkemeier. Passive dynamic quadrupedal walking. In *Proceedings of the 1997 IEEE International Conference on Robotics and Automation*, pp. 34–39. IEEE, 1997.

[77] M. Buehler, A. Cocosco, K. Yamazaki, and R. Battaglia. Stable open loop walking in quadruped robots with stick legs. In *Proceedings of the 1999 IEEE International Conference on Robotics and Automation*, pp. 2348–2353. 1999.

[78] G. Hawker and M. Buehler. Quadruped trotting with passive knees - design, control, and experiments. In *Proceedings of the 2000 IEEE International Conference on Robotics and Automation*. 2000.

[79] Andrew L. Kun and III W. Thomas Miller. Adaptive dynamic balance of an experimental biped robot. *ROBOLEARN 96: An International Workshop on Learning for Autonomous Robots*, pp. 73–78, 1996.

[80] 石黒章夫, 岩木英明, 内川嘉樹. 不整地歩行を可能とする 4 脚歩行ロボットの制御系の進化的 獲得. 第 14 回日本ロボット学会学術講演会講演論文集, pp. 961–962. 日本ロボット学会, 1996.

[81] 橋本亮一, 熊谷徹. 疑似四足歩行ロボットによる適応的歩容獲得実験. 第 14 回日本ロボット学会学術講演会講演論文集, pp. 981–982. 日本ロボット学会, 1996.

[82] John C. Gallagher and Randall D. Beer. A qualitative dynamical analysis of evolved locomotion controllers. In Jean-Arcady Meyer, Herbert L. Roitblat, and Stewart W. Wilson, editors, *From animals to animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, pp. 71–80. The MIT Press, 1993.

[83] 田野浩明, 三上貞芳, 嘉数侑昇. 多足歩行機械の学習による歩容獲得. 第 10 回日本ロボット学会学術講演会予稿集, pp. 1129–1132. 日本ロボット学会, 1992.

[84] 田野浩明, 三上貞芳, 嘉数侑昇. 強化学習を用いた多足歩行機械の適応的歩容計画. ロボティクス・メカトロニクス'93 講演会講演論文集, pp. 863–867. 日本機械学会, 1993.

[85] Pattie Maes and Rodney A. Brooks. Learing to coordinate behaviors. pp. 796–802, 1990.

[86] Kensuke Takita, Keiji Suzuki, and Yukinori Kakazu. Learning mimic impression for multi legged vehicle. In *Proceedings of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 215–222, 1996.

[87] 石黒章夫, 中村弘, 内川嘉樹. 強化学習による 6 脚歩行ロボットの歩容獲得. 第 14 回日本ロボット学会学術講演会講演論文集, pp. 963–964. 日本ロボット学会, 1996.

[88] 近藤啓介, 臼井一哉, 佐竹紀昭, 西川郁子, 渡辺透, 得丸英勝 o. 8 関節 4 足ロボットを用いた遺伝アルゴリズムによる歩行パ ターンの獲得. 第 13 回日本ロボット学会学術講演会予稿集, pp. 221–222. 日本ロボット学会, 1995.

[89] G. S. Hornby, M. Fujita, S. Takamura, T. Yamamoto, and O. Hanagata. Autonomous evolution of gaits with the sony quadruped robot. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 1297–1304. Morgan Kaufmann, 1999.

研究業績

学術雑誌

1. 光永, 浅田. 移動体の意思決定のための情報量基準に基づく観測対象選択戦略. 日本ロボット学会誌, 19(6):793–800, 2001.

2. 光永, 浅田. 情報量に基づく移動ロボットの注視制御のためのセンサ空間構成. 日本ロボット学会誌, 20(7):751–758, 2002.

3. 光永, 浅田. 歩行中の観測を考慮した情報量に基づく注視制御. 日本ロボット学会誌, (投稿中).

国際会議

1. N. Mitsunaga and M. Asada. Observation strategy for decision making based on information criterion. In *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1038–1043. 2000.

2. N. Mitsunaga and M. Asada. Sensor space segmentation for visual attention control of a mobile robot based on information criterion. In *Proceedings of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1714–1719. 2001.

3. N. Mitsunaga and M. Asada. Visual Attention Control for a Legged Mobile Robot based on Information Criterion. In *Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 244–249. 2002.

4. N. Mitsunaga and M. Asada. Observation strategy for decision making based on information criterion. In *Proc. of The Fourth International Workshop on RoboCup*, pp. 32–41, 2000.

5. N. Mitsunaga and M. Asada. Visual attention control by sensor space segmentation for a small quadruped robot based on information criterion. In *The RoboCup 2001 International Symposium*, 2001.

国内口頭発表

1. 光永, 浅田. 移動体の意思決定のための情報量基準による観測戦略. 第5回ロボティクスシンポジア予稿集, pp. 351–356. 2000.
2. 光永, 浅田. 情報量による小型4脚ロボットの行動と観測の効率的決定. 第18回日本ロボット学会学術講演会講演予稿集, pp. 815–816. 日本ロボット学会, 2000.
3. 光永, 浅田. 観測時間と履歴を考慮した情報量に基づく移動ロボットの行動決定のための効率的観測. 第19回日本ロボット学会学術講演会講演予稿集, pp. 659–660. 日本ロボット学会, 2001.
4. 光永, 浅田. 歩行中の観測を考慮した情報量に基づく注視制御. 第20回日本ロボット学会学術講演会講演予稿集, 1B24. 日本ロボット学会, 2002.
5. 光永, 浅田, 野原. エージェントの意思決定のための情報量基準による観測戦略. 人工知能学会 第6回 SIG-Challenge 研究会, pp. 54–59. 2000.
6. 光永, 浅田. 情報量による移動ロボットの注視制御のためのセンサ空間構成. 人工知能学会 第12回 SIG-Challenge 研究会, pp. 29–34. 2001.

紀要等

1. N. Mitsunaga and M. Asada. Observation strategy for decision making based on information criterion. In P. Stone, T. Balch, and G. K. Kraetzschmar eds., *RoboCup 2000: Robot Soccer World Cup IV*, pp. 189–198. Springer, Lecture Note in Artificail Intelligence (2019), 2001.
2. N. Mitsunaga and M. Asada. Visual attention control by sensor space segmentation for a small quadruped robot based on information criterion. In A. Birk, S. Coradeschi, and S. Tadokoro eds., *RoboCup 2001: Robot Soccer World Cup V*, pp. 154–163. Springer, Lecture Note in Artificail Intelligence (2377), 2002.
3. N. Mitsunaga, M. Asada, and C. Mishima. BabyTigers-98: Osaka legged robot team. In M. Asada and H. Kitano eds., *RoboCup-98: Robot Soccer World Cup II*, pp. 498–506. Springer, Lecture Note in Artificail Intelligence (1604), 1999.

4. N. Mitsunaga and M. Asada. BabyTigers-99: Osaka legged robot team. In M. Veloso, E. Pagello, and H. Kitano eds., *RoboCup-99: Robot Soccer World Cup III*, pp. 762–765. Springer, Lecture Note in Artificail Intelligence (1856), 2000.

5. N. Mitsunaga, Y. Nagai, and M. Asada. BabyTigers: Osaka legged robot team. In P. Stone, T. Balch, and G. K. Kraetzschmar eds., *RoboCup 2000: Robot Soccer World Cup IV*, pp. 631–634. Springer, Lecture Note in Artificail Intelligence (2019), 2001.

6. N. Mitsunaga, Y. Nagai, T. Ishida, T. Izumi, and M. Asada. BabyTigers 2001: Osaka legged robot team. In A. Birk, S. Coradeschi, and S. Tadokoro eds., *RoboCup 2001: Robot Soccer World Cup V*, pp. 685–688. Springer, Lecture Note in Artificail Intelligence (2377), 2002.